# Quantile regression methods for first-price auctions

Nathalie Gimenes
Department of Economics
PUC-Rio
Brazil
ngimenes@econ.puc-rio.br

Emmanuel Guerre
School of Economics
University of Kent
United Kingdom
eoguerre@gmail.com

September 2019

## Abstract

The paper proposes a sieve quantile regression approach for first-price auctions with symmetric risk-neutral bidders under the independent private value paradigm. It is first shown that a private value quantile regression model generates a quantile regression for the bids. The private value quantile regression can be easily estimated from the bid quantile regression and its derivative with respect to the quantile level. A new local polynomial technique is proposed to estimate the latter over the whole quantile level interval. Plug in estimation of functionals is also considered, as needed for the expected revenue or the case of CRRA risk-averse bidders, which is amenable to our framework. A quantile regression analysis to USFS timber is found more appropriate than the homogenized bid methodology and illustrates the contribution of each explanatory variables to the private value distribution.

*JEL*: C14, L70

*Keywords*: First-price auction; independent private value; dimension reduction; quantile regression; local polynomial estimation; sieve estimation; boundary correction.

A previous version of this paper has been circulated under the title ”Quantile regression methods for first-price auction:a signal approach”. The authors acknowledge useful discussions and comments from Xiaohong Chen, Valentina Corradi, Yanqin Fan, Phil Haile, Xavier d'Haultfoeuille, Vadim Marmer, Isabelle Perrigne, Martin Pesendorfer and Quang Vuong, and the audience of many conferences and seminars. Nathalie Gimenes also thanks Ying Fan and Ginger Jin for encouragements. All remaining errors are our responsibility. Both authors would like to thank the School of Economics and Finance, Queen Mary University of London, for generous funding.

# 1 Introduction

Various quantile approaches have been recently proposed for the Econometrics of Auctions. Haile, Hong and Shum (2003, HHS hereafter) have used monotonicity of bidding strategy to build a quantile test of the independent private value null hypothesis. Milgrom (2001, Theorem 4.7) reformulates the identification relation of Guerre, Perrigne and Vuong (2000, GPV afterwards) using quantile function. The risk aversion identification result of Guerre, Perrigne and Vuong (2009, GPV09 hereafter) heavily relies on the bid quantile function in first-price auctions. Zincenko (2018) develops a corresponding nonparametric estimation method. Liu and Luo (2017) and Liu and Vuong (2018) have respectively developed quantile based test for the null of exogenous participation and monotonicity of the bidding strategy. Other authors have considered quantile based estimation of the private value distribution. Gimenes (2017) has implemented a quantile regression approach for ascending auction. See also Menzel and Morganti (2013) who proposed an order statistics approach. For first-price auction, Marmer and Shneyerov (2012) has proposed a quantile-based estimator of the private value probability density function (pdf), which is an alternative to the two step GPV method. Guerre and Sabbah (2012) have noted that the private value quantile function can be estimated using a one step procedure from the estimation of the bid quantile function and its first derivative. Enache and Florens (2015) have developed an inverse problem approach.

The two step method of GPV focuses on the private value pdf estimation, which is quite hard to estimate. Estimating pdf is useful for descriptive purposes and for computation of important moments, such as the expected revenue. But the latter can also be achieved using quantile functions, as moments are easily computed integrating it. As noted in Milgrom (2001) in the independent private value setting, the value function of a bidder observing a uniform signal is nothing else than the private value quantile function, so that a quantile approach is especially relevant in auction settings. Nonparametric density estimation is notoriously affected by the curse of dimensionality, and parsimonious models addressing this issue for density are less rich than for quantile functions, where both single index modelling, as already used in an auction framework by Marmer, Shneyerov and Xu (2013b), and additive specification are available. A simpler specification is the homogenized bid model of HHS, which postulates a regression model with iid residuals for the private value. As shown in our empirical application and in Gimenes (2017) for ascending auctions, it may fail to capture nonlinear dependence of the private value to auction covariate. In addition, it still involves a GPV step that may not perform well in small samples.

The present paper develops a quantile regression methodology for first-price auctions, which includes parsimonious but flexible models suitable for moderate samples. The parameter of interest is the private value conditional quantile function given some auction specific covariates, which can be estimated faster than the conditional pdf. A key aspect of our approach is that the bid conditional quantile function is a linear functional of the private value one. It follows that the popular quantile regression model of Koenker and Bassett (1978) can play a central role in our methodology, as it enjoys an important stability property: a private value quantile regression model generates a bid quantile regression model. The private value quantile function is a linear combination of the bid quantile function and its first derivative with respect to the quantile level, a simple identification method which is the basis of our estimation procedure. This also applies to the linear sieve quantile regression of Belloni, Chernozhukov, Chetverikov and Fernández-Val (2017). Following Horowitz and Lee (2005), the latter can be tailored to additive quantile models, which can be better estimated that saturated sieve models. Higher order covariate interactions can also be considered, giving a class of flexible models which can be tailored to each specific datasets.

An important challenge is raised by the estimation of the bid quantile derivative with respect to the quantile level $\alpha$. This was considered by Guerre and Sabbah (2012) and the references therein. We propose instead a new local polynomial approach which applies to quantile levels and aims to jointly estimate the bid quantile function and its derivatives. An unexpected feature is that it performs well for extreme quantile levels, producing consistent estimators for $\alpha = 0$ and 1. The latter upper quantile levels are particularly important for auctions as private values of winners are expected to be in the top of the distribution. Recent work focusing on boundary issues are Aryal, Gabrielli and Vuong (2016) in a semiparametric framework and Hickman and Hubbard (2015). Our theoretical results include a Central Limit Theorem for the private value quantile estimator which holds for extreme quantiles and a bias variance decomposition for its Integrated Mean Squared Error (IMSE). The latter allows in particular for bandwidth choice based on a pilot quantile model.

A second family of parameters of interest consists in integral functionals of the bid quantile function and its quantile level first derivatives. A first example is the parameter of Constant Relative Risk Aversion (CRRA) utility functions. CRRA risk aversion preserves indeed the quantile linearity features which are important for our quantile regression methodology. The risk aversion parameter can be estimated using bidder variations as in GPV09 but also combining first-price and ascending auction as in Lu and Perrigne (2008). A second example is the expected revenue, which falls in such family as it is a functional of the private value quantile function (Gimenes, 2017), see also Li, Perrigne and Vuong (2003). A third example covers the conditional private value cumulative distribution function and pdf. Indeed the rearrangement formula of Chernozhukov, Fernández-Val and Galichon (2010) expresses the cdf as an integral functional of the private value quantile function. Differentiating a smooth version of this functional proposed in Dette and Volgushev (2008) gives a pdf estimator which fits in our framework and differs from Marmer and Shneyerov (2012). These distribution estimators are useful for dimension reduction purpose.

Our theoretical results are illustrated with a simulation experiment and an application to USFS first price auctions. A preliminary quantile regression analysis of the bid quantile function suggests that the homogenized bid technique should not be applied here because the quantile regression slopes are not constant. The private value quantile regression slope functions reveal the impact of the covariate, and how strongly bidders in the top of the distribution can differ from the bottom. CRRA risk-aversion estimation using the approaches of GPV09 and Lu and Perrigne (2008) is also considered. The rest of the paper is organized as follows. The next section introduces our quantile identification approach and the functionals of interest. Section 3 introduces our local polynomial estimation framework. Section 4 groups our main theoretical results for the private value quantile functions and its functionals. Our simulation results are in Section 5 and the application can be found in Section 6. Section 7 summarizes the estimation strategy and the empirical application findings, and describes some possible extensions. All the proofs are gathered in six Online Appendices.

# 2 First price auction and quantile specification

A single and indivisible object with some characteristic $x \in \mathbb{R}^D$ is auctioned to $I \geq 2$ buyers. The potential number of bidders $I$ and $x$ are known to the bidders and the econometrician. Bids are sealed so that a bidder does not know others' bid when forming his own bid. The object is sold to the highest bidder who pays his bid $B_i$ to the seller. Under the symmetric IPV paradigm, each potential bidder is assumed to have a private value $V_i$, $i = 1, \ldots, I$ for the auctioned object. A buyer knows his private value but not the private value of the other bidders, but the joint distribution of the $V_i$ is common knowledge. The private values are independently and identically drawn from a distribution given $(x, I)$ with a compactly supported cdf $F(\cdot|x, I)$, or equivalently with conditional quantile function

$$V(\alpha|x, I) = F^{-1}(\alpha|x, I), \quad \alpha \text{ in } [0, 1].$$

The private value quantile function is the first parameter of interest of the present paper, to be estimated from bids $B_i$ from the symmetric Bayesian Nash equilibrium. Section 2.4 below considers a second set of parameters of interest derived from $V(\cdot|\cdot, \cdot)$ such as the cdf $F(\cdot|\cdot, \cdot)$ or the associated pdf $f(\cdot|\cdot, \cdot)$.

## 2.1 Private value quantile identification

It is well-known that the bidder $i$ private value rank

$$A_i = F(V_i|x, I)$$

has a uniform distribution over $[0,1]$ and is independent of $x$ and $I$. It also follows from the IPV paradigm that the private value ranks $A_i = 1,\ldots,I$ are independent. The dependence between the private value $V_i$ and the auction covariates $x$ and $I$ is therefore fully captured by the non separable quantile representation

$$V_i = V\left(A_i|x,I\right), \quad A_i \overset{\text{iid}}{\sim} \mathcal{U}_{[0,1]} \perp (x,I).$$

Following Milgrom and Weber (1982) or Milgrom (2001), $V(\cdot|x,I)$ can also be viewed as a *valuation function,* the private value rank $A_i$ being the associated signal. In what follows, $G(\cdot|x,I)$ and $g(\cdot|x,I)$ stand for respectively the bid conditional cdf and pdf.

Maskin and Riley (1984) have shown that Bayesian Nash Equilibrium bids $B_i = \sigma(V_i;x,I)$ of symmetric risk averse or risk neutral bidders must strictly and continuously increase with the private values under the IPV paradigm. It follows that $B_i = B(A_i|x,i)$ where $B(\cdot;x,i) = \sigma(F(\cdot|x,I);x,I)$ can be viewed as a bidding strategy depending upon the rank $A_i$. If $F(\cdot|x,I)$ is also strictly increasing, so is $B(\cdot|x,I)$ and since $A_i$ is uniform it holds

$$G(b|x,I) = \mathbb{P}\left[B(A_i|x,I) \leq b|x,I\right] = \mathbb{P}\left[A_i \leq B^{-1}(b|x,I)\,|x,I\right] = B^{-1}(b|x,I)$$

showing that the bidding strategy $B(\cdot|x,I)$ is also the bid quantile function.

A standard best response argument will show how to identify the private value quantile function $V(\cdot|x,I)$ from $B(\cdot|x,I)$. Suppose bidder $i$ signal $A_i$ is equal to $\alpha$, but that her bid is a suboptimal $B(a|x,I)$, all other bidders bidding $B(A_j|x,I)$. Then the probability that bidder $i$ wins the auction is

$$\begin{aligned}\mathbb{P}\left[B(a|x,I) > \max_{1\leq j\neq i\leq I} B(A_j|x,I)\middle| A_i=\alpha,x,I\right] &= \mathbb{P}\left[a > \max_{1\leq j\neq i\leq I} A_j \middle| A_i=\alpha,x,I\right]\\ &= a^{I-1}\end{aligned} \tag{2.1}$$

because the $A_j$'s are independent $\mathcal{U}_{[0,1]}$ independent of $x$ and $I$. It follows that the expected revenue of such a bid is, for a risk neutral bidder, $(V(\alpha|x,I) - B(a|x,I))\,a^{I-1}$. If $B(\cdot|x,I)$

is a best-response bidding strategy, the optimal bid of a bidder with signal $\alpha$ is $B(\alpha|x, I)$, that is

$$\alpha = \arg\max_{a} \left\{ \left( V(\alpha|x, I) - B(a|x, I) \right) a^{I-1} \right\}.$$

As $B(\cdot|x, I)$ is continuously differentiable, it follows that

$$\frac{\partial}{\partial a} \left\{ \left( V(\alpha|x, I) - B(a|x, I) \right) a^{I-1} \right\} \Big|_{a=\alpha} = 0 \tag{2.2}$$

or equivalently $\frac{d}{d\alpha} \left[ \alpha^{I-1} B(\alpha|x, I) \right] = (I-1) \alpha^{I-2} V(\alpha|x, I)$. Solving with the initial condition $B(0|x, I) = V(0|x, I)$ and rearranging the equation above gives Proposition 1, which is the cornerstone of our estimation method. From now on $B^{(1)}(\alpha|x, I) = \frac{d}{d\alpha} B(\alpha|x, I)$.

**Proposition 1** *Consider a given $(x, I)$, $I \geq 2$, for which $\alpha \in [0, 1] \mapsto V(\alpha|x, I)$ is continuously differentiable with a derivative $V^{(1)}(\cdot|x, I) > 0$. Suppose the bids are drawn from the symmetric differential Bayesian Nash equilibrium. Then,*

*i. The conditional equilibrium quantile function $B(\cdot|x, I)$ of the $I$ iid optimal bids $B_i$ satisfies*

$$B(\alpha|x, I) = \frac{I-1}{\alpha^{I-1}} \int_0^{\alpha} a^{I-2} V(a|x, I)\, da. \tag{2.3}$$

*ii. The bid quantile function $B(\alpha|x, I)$ is continuously differentiable over $[0, 1]$ and it holds*

$$V(\alpha|x, I) = B(\alpha|x, I) + \frac{\alpha B^{(1)}(\alpha|x, I)}{I-1}. \tag{2.4}$$

A key feature is the linearity of the private value to bid quantile function mapping (2.3), which implies that a private value quantile linear model is mapped into a similar bid linear model, as detailed below for the well known quantile regression. Proposition 1-(ii) shows that the private value quantile function is identified from the bid quantile function and its derivative, as noted in Guerre and Sabbah (2012). It is a quantile version of the identification strategy of GPV, based on the computation of the private value from the bid[1]

$$V_i = B_i + \frac{1}{I-1} \frac{G(B_i|x, I)}{g(B_i|x, I)}.$$

Versions of (2.4) with $B^{(1)}(\alpha|x, I)$ changed into $1/g(B(\alpha|x, I)|x, I)$ can be found in Milgrom (2001, Theorem 4.7), Liu and Luo (2014), Enache and Florens (2015), Liu and Vuong (2016) and Luo and Wan (2016) and, under risk aversion, in GPV09 and Campo, Guerre, Perrigne and Vuong (2011). As developed in Section 2.4 below, Proposition 1 can be extended to the case of symmetric risk-averse bidders with a CRRA utility function.

## 2.2 Private value quantile regression and homogenized bids

**Private value quantile regression.** The linearity of (2.3) with respect to the private value quantile function has remained unnoticed with very few exceptions, although it has important model stability implications useful for practical implementation. Consider for instance a private value quantile given by the quantile regression specification

$$V(\alpha|x, I) = \gamma_0(\alpha|I) + x'\gamma_1(\alpha|I) = [1, x']\gamma(\alpha|I). \tag{2.5}$$

Proposition 1-(i) implies that the conditional bid quantile function satisfies,

$$B(\alpha|x, I) = [1, x']\beta(\alpha|I) \text{ with } \beta(\alpha|I) = \frac{I-1}{\alpha^{I-1}} \int_0^\alpha t^{I-2}\gamma(t|I)\,dt, \tag{2.6}$$

showing $B(\alpha|x, I)$ belongs to the quantile regression specification. It follows from (2.4) that

$$\gamma(\alpha|I) = \beta(\alpha|I) + \frac{\alpha\beta^{(1)}(\alpha|I)}{I-1}, \tag{2.7}$$

[1]This can be recovered from (2.4) taking $\alpha = A_i$ as $V_i = V(A_i|x, I)$, $B_i = B(A_i|x, I)$ implying that $A_i = G(A_i|x, I)$ and $B^{(1)}(A_i|x, I) = 1/g(B(A_i|x, I)|x, I) = 1/g(B_i|x, I)$.

so that $\gamma(\alpha|I)$ can easily be estimated from an estimation of $\beta(\alpha|I)$ and $\beta^{(1)}(\alpha|I)$. It then follows that the quantile regression specification is stable, i.e. a quantile regression specification for the private value is equivalent to a quantile regression specification for the bid. Hence testing the correct specification of a bid quantile regression model is equivalent to test the correct specification of a private value quantile specification. The expressions (2.6) and (2.7) show that significance testing can be done through bid quantile regression as $\gamma_j(\cdot|I) = 0$ is equivalent to $\beta_j(\cdot|I) = 0$, or more generally $e'\gamma(\cdot|I) = c$ is equivalent to $e'\beta(\cdot|I) = c$ for any conformable $e$ and $c$.

**Bid homogenization and quantile regression.** HHS have noted that a translation of the private values results in a similar translation of the bids, an invariance property that they use in their bid homogenization technique. The latter can be interpreted as the use of a regression model for the private values, $V_i = \gamma_0 + x'\gamma_1 + v_i$ with an error term $v_i$ independent of $x$, as also proposed by Rezende (2008). This amounts to assume that the slope function $\gamma_1(\cdot|I)$ in (2.5) does not depend upon the quantile level. The regression model of HHS and Rezende (2008) is indeed equivalent to the quantile regression specification

$$V(\alpha|x) = \gamma_0 + x'\gamma_1 + v(\alpha)$$

where $v(\alpha)$ is the quantile function of $v_i$. Since $\frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2}da = 1$, it follows that the associated bid quantile function is, by (2.3)

$$B(\alpha|x, I) = \gamma_0 + x'\gamma_1 + b(\alpha|I), \text{ where } b(\alpha|I) = \frac{I-1}{\alpha^{I-1}}\int_0^\alpha a^{I-2}v(a)\,da.$$

This gives the bid regression model

$$B_i = \beta_0(I) + x'\gamma_1 + b_i, \quad \beta_0(I) = \gamma_0 + \mathbb{E}\left[b(A_i|I)\right]$$

where the regression error term $b_i = b(A_i|I) - \mathbb{E}\left[b(A_i|I)\right]$ is centered and independent of $x$. Following these authors, the coefficient $\gamma_1$ can be estimated regressing the bids on $[1, x']$

and the distribution of $v_i$ can be estimated applying the GPV two step method to the homogenized bids, which are the residuals $B_i - x'\widehat{\gamma}_1$.

However this approach requests independence between the regression error term $v_i$ and the covariate $x$, an assumption which may be too restrictive in practice as found by Gimenes (2017) and the application below. When $\gamma_1(\cdot)$ is not a constant, regressing $B(\alpha|x, I)$ on $[1, x]$ gives $B_i = \beta_0(I) + x'\beta_1(I) + b(A_i|x, I)$ with a slope coefficient satisfying

$$\begin{aligned}\beta_1(I) &= \int_0^1 \left( \frac{I-1}{\alpha^{I-1}} \int_0^\alpha a^{I-2} \gamma_1(a)\, da \right) d\alpha \\ &= \int_0^1 \gamma_1(\alpha)\, d\alpha - \int_0^1 \left( \int_0^\alpha \left(\frac{a}{\alpha}\right)^{I-1} \gamma_1^{(1)}(a)\, da \right) d\alpha\end{aligned}$$

and a residual term $b(A_i|x, I) = v(A_i) + x' \frac{I-1}{A_i^{I-1}} \int_0^{A_i} a^{I-2} \gamma_1(a)\, da - \beta_1(I)$ which now depends upon $x$, so that the homogenized bid approach does not apply. Using variation of $I$ can be useful to detect such a situation because observing variation of $\beta_1(I)$ implies that $\gamma_1(\cdot)$ is not a constant. In particular, If the entries of $\gamma_1^{(1)}(\cdot)$ are nonnegative, the entries of $\beta_1(I)$ must increase with $I$. Similar features hold for centered bids $B_i - \mathbb{E}[B_i|I]$ when the homogenized bid regression is replaced by a nonparametric regression: the regression function $\mathbb{E}[B_i - \mathbb{E}[B_i|I]|\, x, I]$ should not depend upon $I$ if $V_i = m(X) + v_i$, as for the single index regression specification considered in Paarsch and Hong (2006).

## 2.3 Linear nonparametric quantile specification

**Flexible interactive specifications.** The private value quantile regression model (2.5) assumes linearity of the private value quantile function with respect to the covariate $x$. This may be too strong and can be relaxed using a quantile nonparametric additive specification, which was considered in Horowitz and Lee (2005). Recall that $x = (x_1, \ldots, x_D)$ and consider the additive quantile function

$$V(\alpha|x, I) = \sum_{j=1}^{D} V_j(\alpha; x_j, I) \tag{2.8}$$

where each functions $V_j(\alpha; x_j, I)$ is specific to the entry $x_j$. Since such quantile specifications are obtained by summing some univariate functions, the effective dimension involved in the nonparametric dimension of this model is 1 because it can be estimated with the same rate than a nonparametric model with a unique covariate as shown in Horowitz and Lee (2005). This parsimonious model can be generalized following Andrews and Whang (1990) to allow for more covariate interactions. This leads to the additive interactive quantile specification with $D_{\mathcal{M}}$ interactions

$$V(\alpha|x, I) = \sum_{\delta=1}^{D_{\mathcal{M}}} \sum_{1 \leq j_1 < \cdots < j_\delta \leq d} V_{j_1 \ldots j_\delta}(\alpha; x_{j_1}, \ldots x_{j_\delta}, I) \tag{2.9}$$

where each functions $V_{j_1 \cdots j_\delta}(\alpha; x_{j_1}, \ldots x_{j_\delta}, I)$ can now depend upon $\delta$ entries of $x$ with $\delta \leq D_{\mathcal{M}} \leq D$. Setting $D_{\mathcal{M}}$ equal to the dimension $D$ of the covariate gives the general quantile specification. As seen from Andrews and Whang (1990) for the regression case, such specification can be estimated with the same rate than a function of $D_{\mathcal{M}}$ variables, so that $D_{\mathcal{M}}$ can be viewed as the effective dimension of this model.

The stability property in Proposition 1-(i) ensures that a private value quantile specification with $D_{\mathcal{M}}$ interaction will generate a bid quantile specification with the same number of interactions: if (2.9) holds, then the bid quantile function satisfies

$$B(\alpha|x, I) = \sum_{\delta=1}^{D_{\mathcal{M}}} \sum_{1 \leq j_1 < \cdots < j_\delta \leq d} B_{j_1 \ldots j_\delta}(\alpha; x_{j_1}, \ldots x_{j_\delta}, I)$$

and the private values components of the specification can be recovered using Proposition 1-(ii).

**Sieve interactive specification.** The interactive quantile specification (2.9) can be estimated using a sieve expansion, as in Horowitz and Lee (2005) or Andrews and Whang (1990). Consider a sieve $\{P_k(x), 1 \leq k \leq K\}$ is a family of functions $P_k(\cdot) = P_{kK}(\cdot)$ allowing for at most $D_{\mathcal{M}}$ interactions and suppose that there are some sieve coefficients $\gamma_k(\cdot|I) = \gamma_{kK}(\cdot|I)$

such that for all $\alpha$

$$V(\alpha|x,I)=\lim_{K\to\infty}\sum_{k=1}^{K}\gamma_k(\alpha|I)\,P_k(x)\,. \tag{2.10}$$

The expression (2.10) can be viewed as a sieve extension of the quantile regression, a *sieve quantile regression*. It follows from Proposition 1-(i,ii) that, provided the limit in (2.10) holds uniformly with respect to $\alpha$,

$$B(\alpha|x,I)=\lim_{K\to\infty}\sum_{k=1}^{K}\beta_k(\alpha|I)\,P_k(x)\,,\ \ \beta_k(\alpha|I)=\frac{I-1}{\alpha^{I-1}}\int_0^{\alpha}t^{I-2}\gamma_k(t|I)\,dt, \tag{2.11}$$

$$V(\alpha|x,I)=\lim_{K\to\infty}\sum_{k=1}^{K}\left(\beta_k(\alpha|I)+\frac{\alpha\beta_k^{(1)}(\alpha|I)}{I-1}\right)P_k(x)\,. \tag{2.12}$$

Hence estimating the private value sieve quantile regression can proceed from estimating the coefficients of the bid sieve quantile regression in (2.11) and their first derivatives.

## 2.4 Risk aversion, expected payoff and other functionals

Many auction parameters of interest can be written using the private value quantile functions, or equivalently the bid quantile function and its quantile derivative by (2.4). We focus here on the conditional and unconditional integral functionals

$$\theta(x)=\int_0^1\mathcal{F}\left[\alpha,x,B(\alpha|x,I)\,,B^{(1)}(\alpha|x,I)\,;I\in\mathcal{I}\right]d\alpha,\quad \theta=\int_{\mathcal{X}}\theta(x)\,dx \tag{2.13}$$

where $\mathcal{F}(\alpha,x,b_{0I},b_{1I};I\in\mathcal{I})$ is a real valued continuous function. Three illustrative examples are as follows.

**Example 1: CRRA risk aversion.** For symmetric risk averse bidders with a concave utility function, the best response condition (2.2) becomes

$$\frac{\partial}{\partial a}\left\{U\left(V(\alpha|x,I)-B(a|x,I)\right)a^{I-1}\right\}\Big|_{a=\alpha}=0.$$

Rearranging as in GPV09 yields that $V(\alpha|x,I) = B(\alpha|x,I) + \lambda^{-1}\left(\frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}\right)$ where $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. For risk averse bidders with a CRRA utility function $U(t) = t^{\theta}$, arguing as for Proposition 1 shows

$$\begin{aligned} V(\alpha|x,I) &= B(\alpha|x,I) + \theta\frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}, \\ B(\alpha|x,I) &= \frac{1}{\alpha^{\frac{I-1}{\theta}}}\int_0^{\alpha} t^{\frac{I-1}{\theta}-1} V(t|x,I)\,dt. \end{aligned} \tag{2.14}$$

These two formulas show that the stability implications of Proposition 1 for linear private value and bid quantile functions are preserved under CRRA. Assuming as in GPV09 that the number of bidders is exogenous, i.e $V(\alpha|x,I) = V(\alpha|x)$ for all $I$, gives, for any pair $I_0 \neq I_1$

$$\theta = \frac{\theta_n}{\theta_d} = \frac{\int_{\mathcal{X}}\left[\int_0^1 \left(B(\alpha|x,I_1) - B(\alpha|x,I_0)\right)\left(\frac{\alpha B^{(1)}(\alpha|x,I_0)}{I_0-1} - \frac{\alpha B^{(1)}(\alpha|x,I_1)}{I_1-1}\right)d\alpha\right]dx}{\int_{\mathcal{X}}\left[\int_0^1 \left(\frac{\alpha B^{(1)}(\alpha|x,I_0)}{I_0-1} - \frac{\alpha B^{(1)}(\alpha|x,I_1)}{I_1-1}\right)^2 d\alpha\right]dx}, \tag{2.15}$$

a formula which shows that the CRRA risk aversion can be easily identified from first-price auction. Following Lu and Perrigne (2008), the risk-aversion parameter $\theta$ can also be identified combining ascending and first-price auctions data. As seen from Gimenes (2017), the private value quantile function $V_{asc}(\alpha|x,I)$ can be easily estimated from ascending auctions. Equating $V_{asc}(\alpha|x,I)$ to $V(\alpha|x,I)$ in (2.14) gives that $\theta$ satisfies

$$\theta = \frac{\int_{\mathcal{X}}\left[\int_0^1 \left(V_{asc}(\alpha|x,I) - B(\alpha|x,I)\right)\frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}d\alpha\right]dx}{\int_{\mathcal{X}}\left[\int_0^1 \left(\frac{\alpha B^{(1)}(\alpha|x,I)}{I-1}\right)^2 d\alpha\right]dx}. \tag{2.16}$$

**Example 2: Expected revenue.** Suppose that the seller decides to reject bids lower than a reserve price $R$ and let $\alpha_R = \alpha_R(x,I)$ be the associated screening level, i.e. $\alpha_R =$

$F(R|x,I)$. For CRRA bidders, the first price auction seller's expected revenue is[2]

$$\begin{aligned} ER_\theta(\alpha_R|x,I) = & \frac{\theta \cdot I \cdot V(\alpha_R|x,I)}{(I-1)(\theta-1)+\theta}\alpha_R^{\frac{I-1}{\theta}}\left(1-\alpha_R^{(I-1)\frac{\theta-1}{\theta}+1}\right) \\ & + \frac{I(I-1)}{(I-1)(\theta-1)+\theta}\int_{\alpha_R}^1 t^{\frac{I-1}{\theta}-1}\left(1-t^{(I-1)\frac{\theta-1}{\theta}+1}\right)V(t|x,I)\,dt. \end{aligned} \tag{2.17}$$

This expression includes an integral item

$$\theta(x;\alpha_R) = \int_{\alpha_R}^1 t^{\frac{I-1}{\theta}-1}\left(1-t^{(I-1)\frac{\theta-1}{\theta}+1}\right)V(t|x,I)\,dt$$

which can be estimated by plugging in a risk aversion estimator $\widehat{\theta}$ and an estimator $\widehat{V}(\alpha|x,I)$ of the private value quantile function, or estimators of the bid quantile function and its derivative by (2.4).[3]

**Example 3: Private value distribution** Chernozhukov et al. (2010) have used the rearrangement formula to invert a monotonic function. In our case, the conditional private value cdf satisfies

$$F(v|x,I) = \mathbb{E}\left[\mathbb{I}\left[V(A|x,I) \le v\right]|\,x,I\right] = \int_0^1 \mathbb{I}\left[V(\alpha|x,I) \le v\right]d\alpha, \quad A \sim \mathcal{U}_{[0,1]}.$$

Dette and Volgushev (2008) have considered a smoothed version $\mathbb{I}_\eta(\cdot)$ of the indicator function

$$F_\eta(v|x,I) = \int_0^1 \mathbb{I}_\eta\left[v - V(\alpha|x,I)\right]d\alpha$$

[2]It is assumed for the sake of brevity that the seller value for the good is 0.The expected revenue formula for the general case follows from Gimenes (2017).

[3]Under risk-neutrality, integrating by parts gives that

$$\int_{\alpha_R}^1 B^{(1)}(\alpha|x,I)\,\alpha^{I-1}(1-\alpha)\,d\alpha = B(\alpha_R|\,x,I)\,\alpha_R^{I-1}(1-\alpha_R) - \int_{\alpha_R}^1 B(\alpha|x,I)\,\alpha^{I-1}(I-1-I\alpha)\,d\alpha,$$

estimation of $\theta(x;\alpha_R)$ can also be done using only a bid quantile estimator.

where $\mathbb{I}_{\eta}(t) = \int_{-\infty}^{t/\eta} k(u)\, du$, $k(\cdot)$ being a kernel function and $\eta$ a bandwidth parameter. Differentiating $F_{\eta}(v|x,I)$ gives

$$f_{\eta}(v|x,I) = \frac{1}{\eta}\int_{0}^{1} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right) d\alpha$$

which converges to the private value pdf when $\eta$ goes to 0. Note that $F_{\eta}(v|x,I)$ and $f_{\eta}(v|x,I)$ can be estimated by plugging in an estimator $\widehat{V}(\alpha|x,I)$ of $V(\alpha|x,I)$. The resulting cdf and pdf estimator are expected to inherit of the dimension reduction property of this procedure. As the private value estimator $\widehat{V}(\alpha|x,I)$ proposed in the next section is consistent over the whole $[0,1]$, no trimming is needed. This contrasts with the GPV pdf estimator.

# 3 Augmented quantile regression estimation

Proposition 1 suggests to base the estimation of the private value quantile function on estimations of $B(\alpha|x,I)$ and of its derivative $B^{(1)}(\alpha|x,I)$ with respect to $\alpha$. While there is an important literature on the estimation of a conditional quantile function, estimating the first derivative of a quantile function has received much less attention. The *augmented* methodology applies local polynomial expansion with respect to $\alpha$ for joint estimation of $B(\alpha|x,I)$ and $B^{(1)}(\alpha|x,I)$. Sieve methods can be used for the covariate. To ensure comparability with the literature, we assume that the private value quantile function $V(\alpha|x,I)$ has $s+1$ continuous derivatives with respect to $\alpha$. As seen from (2.3), this implies that the bid quantile function $B(\alpha|x,I)$ has $s+2$ continuous derivatives with respect to $\alpha > 0$. This justifies the order $s+1$ for the local polynomial estimator considered here.

## 3.1 Definition of the estimators

**The no covariate case.** Consider $L$ iid first-price auctions $(I_{\ell}, x_{\ell}, B_{i\ell}, i = 1, \ldots, I_{\ell})$. To introduce our estimation strategy, assume first that $V(\alpha|x,I) = V(\alpha|I)$ and $B(\alpha|x,I) =$

$B(\alpha|I)$. Let $\rho_\alpha(\cdot)$ be the check function,

$$\rho_\alpha(q) = q(\alpha - \mathbb{I}(q \le 0)),$$

$\mathbb{I}(\cdot)$ being the indicator function, $\mathbb{I}(q \le 0) = 1$ for $q \le 0$ and 0 otherwise. It is well known that,

$$B(\alpha|I) = \arg\min_q \mathbb{E}[\mathbb{I}(I_\ell = I)\rho_\alpha(B_{i\ell} - q)], \quad \alpha \in (0,1).$$

Estimating the derivative $B^{(1)}(\alpha|I)$ can be done by introducing local variation of the quantile level in the vicinity of $\alpha$. Let $K(\cdot) \ge 0$ be a kernel function with support $[-1,1]$ and $h = h_L$ be a positive bandwidth parameter going to 0 with the sample size. Then it follows that

$$\begin{aligned} &\{B(a|I), a \in [\alpha - h, \alpha + h] \cap [0,1]\} \\ &= \arg\min_{q(a)} \int_0^1 \mathbb{E}[\mathbb{I}(I_\ell = I)\rho_a(B_{i\ell} - q(a))] \frac{1}{h} K\left(\frac{a-\alpha}{h}\right) da, \end{aligned} \tag{3.1}$$

where the minimization is performed over the set of functions $q(a)$ which are continuous on $[\alpha - h, \alpha + h] \cap [0,1]$. Instead of a minimization over such a rich set of functions, we consider minimization over a set of polynomial functions. Indeed, a good polynomial approximation of $B(a|I)$ over $[\alpha - h, \alpha + h]$ is given by the Taylor expansion

$$B(a|I) = B(\alpha|I) + B^{(1)}(\alpha|I)(a-\alpha) + \cdots + \frac{B^{(s+1)}(\alpha|I)(a-\alpha)^{s+1}}{(s+1)!} + O\left(h^{s+2}\right).$$

Let $b = (\beta_0, \ldots, \beta_{s+1})'$ be the generic coefficients of such a polynomial function and

$$\pi(a) = \left[1, a, \frac{a^2}{2} \ldots, \frac{a^{s+1}}{(s+1)!}\right]'.$$

The sample version of the objective function (3.1) restricted to polynomial functions is

$$\begin{aligned}\widehat{\mathcal{R}}\left(b;\alpha,I\right) &= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_\ell=I\right)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell}-\pi\left(a-\alpha\right)'b\right)\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da\\ &= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_\ell=I\right)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell}-\pi\left(ht\right)'b\right)K\left(t\right)dt.\end{aligned}$$

The *augmented quantile* estimator is $\widehat{b}\left(\alpha|I\right)=\arg\min_b\widehat{\mathcal{R}}\left(b;\alpha,I\right)$, $\widehat{\beta}_0\left(\alpha|I\right)$ and $\widehat{\beta}_1\left(\alpha|I\right)$ being estimators of $B\left(\alpha|I\right)$ and its first derivative $B^{(1)}\left(\alpha|I\right)$, respectively.[4] The estimator of the private value quantile is[5]

$$\widehat{V}\left(\alpha|I\right)=\widehat{\beta}_0\left(\alpha|I\right)+\frac{\alpha\widehat{\beta}_1\left(\alpha|I\right)}{I-1}.$$

**Augmented quantile regression.** A first extension of this procedure is the *augmented quantile regression* estimator, AQR hereafter, which considers the private quantile regression specification

$$V\left(\alpha|x,I\right)=\left[1,x'\right]\gamma\left(\alpha|I\right).$$

[4]When the private value distribution does not depend upon $I$, the bid quantile functions $B\left(\cdot|I\right)$ are such that the derivatives

$$\frac{\partial^j}{\partial\alpha^j}\left[B\left(\alpha|I\right)+\frac{\alpha B^{(1)}\left(\alpha|I\right)}{I-1}\right]=\left(1+\frac{j}{I-1}\right)B^{(j)}\left(\alpha|I\right)+\frac{\alpha B^{(j+1)}\left(\alpha|I\right)}{I-1}$$

do not depend upon $I$ as they are equal to $V^{(j)}\left(\alpha|I\right)=V^{(j)}\left(\alpha\right)$, $j=0,\ldots,s+1$. These constraints can be used to estimate $V\left(\alpha\right)$ using the parameters $\gamma=\left(\gamma_0,\ldots\gamma_s\right)$, $\delta=\left(\delta_2,\ldots,\delta_{\overline{I}}\right)$ where $\gamma_j$ is for $V^{(j)}\left(\alpha\right)$ and $\delta_I$ for the derivatives $B^{(s+1)}\left(\alpha|I\right)$, $I=2,\ldots,\overline{I}$ and $b_I\left(\gamma,\delta\right)=\left[b_{0,I},\ldots,b_{s,I},\delta_I\right]'$ with $b_{s,I}=\left(1+\frac{s}{I-1}\right)^{-1}\left(\gamma_s-\frac{\alpha}{I-1}\delta_I\right)$ and the $b_{j,I}$'s are computed recursively using

$$b_{j,I}=\left(1+\frac{j}{I-1}\right)^{-1}\left(\gamma_j-\frac{\alpha}{I-1}b_{j+1,I}\right),\ j=0,\ldots,s.$$

The estimator of $V\left(\alpha\right)$ is $\widehat{\gamma}_0$ where $\left(\widehat{\gamma},\widehat{\delta}\right)=\arg\min_{\gamma,\delta}\sum_{I=2}^{\overline{I}}\widehat{\mathcal{R}}\left(b_I\left(\gamma,\delta\right);\alpha,I\right)$.

[5]Although not considered here, the augmented quantile estimation procedure can be used to estimate the p.d.f. $f\left(v|I\right)$ of the private value using $f\left(v|I\right)=1/V^{(1)}\left[F\left(v|I\right)|I\right]$. An estimator for $F\left(\cdot|I\right)$ is $\widehat{V}^{-1}\left(\cdot|I\right)$. Set $\widehat{V}^{(1)}\left(\alpha|I\right)=\widehat{\beta}_1\left(\alpha|I\right)+\alpha\widehat{\beta}_2\left(\alpha|I\right)/\left(I-1\right)$ and $\widehat{f}\left(v|I\right)=1/\widehat{V}^{(1)}\left[\widehat{F}\left(v|I\right)|I\right]$. This p.d.f. estimator can account for covariates by using the AQR and ASQR procedures introduced below.

In the second extension, the *augmented sieve quantile regression* (ASQR), the private value quantile function $V(\alpha|x,I)$ is equal to $P(x)'\gamma(\alpha|I)$ up to an approximation error, where $P(x)$ stacks the sieve functions $P_k(x)$, $k=1,\ldots,K$. The AQR and ASQR approaches can be grouped setting $P(x)=[1,x']'$ for the AQR.

The bid quantile function satisfies $B(\alpha|x,I) = P(x)'\beta(\alpha|I)$ by (2.6) with $\gamma(\alpha|I) = \beta(\alpha|I) + \alpha\beta^{(1)}(\alpha|I)/(I-1)$ by (2.7), up to an approximation error in the ASQR case. Define now the parameter

$$b = \left[\beta_0', \beta_1', \ldots, \beta_{s+1}'\right]$$

where all the $\beta_j$ have the same dimension $D+1$ and

$$P(x,t) = \pi(t) \otimes P(x)$$

which is such that the Taylor expansion of $B(\alpha|x,I)$ writes, in the AQR case,

$$B(\alpha+ht|x,I) = P(x,ht)'b(\alpha|I) + O\left(h^{s+2}\right)$$

where $b(\alpha|I)$ stacks $\beta(\alpha|I)$ and its successive derivatives $\beta^{(1)}(\alpha|I),\ldots,\beta^{(s+1)}(\alpha|I)$. The objective function of the estimation procedure becomes

$$\begin{aligned}
\widehat{\mathcal{R}}(b;\alpha,I) &= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell=I)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell}-P(x_\ell,a-\alpha)'b\right)\frac{1}{h}K\left(\frac{a-\alpha}{h}\right)da \\
&= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell=I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \rho_{\alpha+ht}\left(B_{i\ell}-P(x_\ell,ht)'b\right)K(t)\,da \qquad (3.2)
\end{aligned}$$

which accounts for the covariate $x_\ell$. The estimation of $b(\alpha|I)$ is $\widehat{b}(\alpha|I) = \arg\min_b \widehat{\mathcal{R}}(b;\alpha,I)$ and the private value quantile regression estimator is

$$\widehat{V}(\alpha|x,I) = P(x)'\widehat{\gamma}(\alpha|I) \text{ with } \widehat{\gamma}(\alpha|I) = \widehat{\beta}_0(\alpha|I) + \frac{\alpha\widehat{\beta}_1(\alpha|I)}{I-1}.$$

The bid quantile function and its derivatives can be estimated using $\widehat{B}(\alpha|x,I) = P(x)'\widehat{\beta}_0(\alpha|I)$

and $\widehat{B}^{(1)}(\alpha|x,I) = P(x)'\widehat{\beta}_1(\alpha|I)$. The rearrangement method of Chernozhukov et al. (2010) can be used to obtain increasing quantile estimators.

## 3.2 Boundary estimation

Bassett and Koenker (1982) report that standard quantile regression estimators are not defined for the extreme quantile levels $\alpha = 0$ or $\alpha = 1$ or even nearby. The augmented procedures proposed here are better behaved for extreme quantiles because the objective function $\widehat{\mathcal{R}}(\cdot;\alpha,I)$ averages the check function $\rho_a(\cdot)$ for quantile levels $a$ in $[\alpha-h,\alpha+h]\cap[0,1]$. For instance, if $\alpha = 1$ and $h \leq 1$, $\widehat{\mathcal{R}}(b;1,I)$ averages $\rho_{1+ht}\left(B_{i\ell} - P(x_\ell,ht)'b\right)$ over $t$ in $[-1,0]$, so that $\widehat{\mathcal{R}}(b;1,I)$ will be large if $b$ is too large.[6] Figure 1 below shows indeed that $\widehat{\mathcal{R}}(b;1,I)$ has no flat part when $b$ grows, contrasting with the standard quantile regression objective functions.

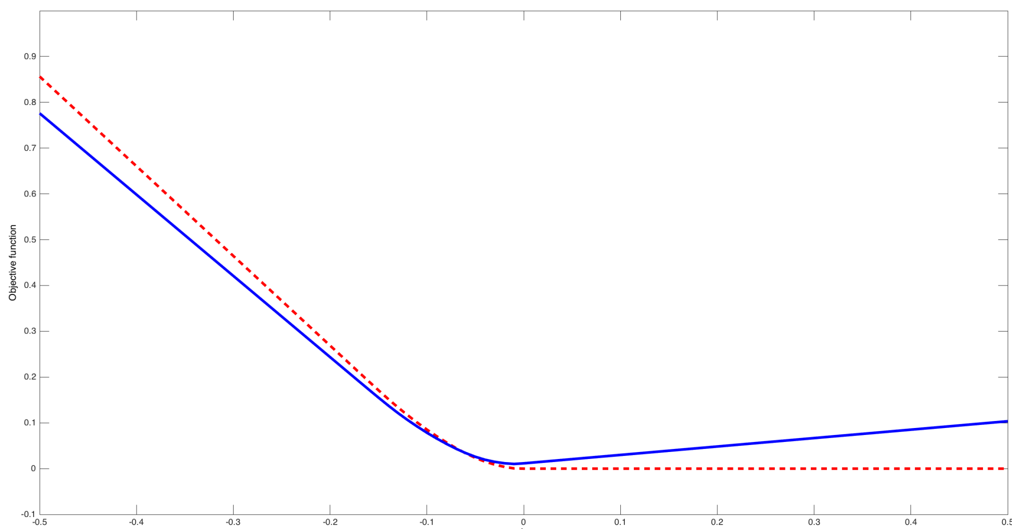


Figure 1: A path of the objective function $\widehat{\mathcal{R}}(b\cdot;1,I)$ (solid line) of the augmented quantile regression estimator and of the objective function of the standard quantile regression estimator (dotted line) when $b$ varies in the direction $[1,\ldots,1]'$.

[6]This averaging effect requests that $t \to P(x_\ell,ht)'b$ is not constant meaning that the derivative components of $b$ should not vanish.

Therefore the AQR and ASQR estimators are easier to define for the extreme quantile levels $\alpha = 0$ and $\alpha = 1$ than the standard quantile regression estimator. This is especially relevant for estimating auction models as the winner is expected to belong to the upper tail as soon as the number of bidders is large enough. In fact, it follows from the theoretical study of the objective function $\widehat{\mathcal{R}}(\cdot;\cdot, I)$ that the AQR and ASQR estimators are uniquely defined for all quantile levels with a large probability.[7] As a result of a smooth objective function, the AQR and ASQR estimators are also smoother than standard quantile regression ones, see for instance Figure 4 in the Application Section.

# 4 Main results

## 4.1 Main assumptions and sieve choice

The notations $a \vee b$ and $a \wedge b$ are used instead of $\max(a, b)$ and $\min(a, b)$. Recall $a_L \asymp b_L$ means that both $a_L/b_L = O(1)$ and $b_L/a_L = O(1)$. The norm $\|\cdot\|$ is the Euclidean one, i.e. $\|e\| = (e'e)^{1/2}$.

### 4.1.1 General assumptions

**Assumption A** *(i) The auction variables $(I_\ell, x_\ell, V_{i\ell}, B_{i\ell}, i = 1, \ldots, I_\ell)$ are iid across $\ell$. The pdf $f(x|I)$ of the covariates $x_\ell$ given $I_\ell = I$ is continuous and bounded away from 0 over its bounded support $\mathcal{X}$, with a non empty interior and which does not depend upon $I$. The actual number of bidders $I_\ell$ belongs to a finite set $\mathcal{I}$ of integer numbers larger or equal to 2.*

*(ii) Given $(x_\ell, I_\ell) = (x, I)$, the $V_{i\ell}$, $i = 1, \ldots, I_\ell$ are iid with a conditional quantile function $V(\alpha|x, I)$, which is continuously differentiable over $[0,1] \times \mathcal{X}$ with*

$$\inf_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) > 0 \text{ and } \sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}} V^{(1)}(\alpha|x, I) < \infty.$$

*(iii) (2.3) holds with $B(0|x, I) = V(0|x, I)$ for all $(x, I) \in \mathcal{X} \times \mathcal{I}$.*

[7] See the discussion following Theorem C.4 in Appendix C for a formal argument.

**Assumption S** *For some* $s \geq 1$ *and each* $I \in \mathcal{I}$, $V(\alpha|x, I)$ *is* $(s+1)-$*times continuously differentiable over* $[0,1] \times \mathcal{X}$ *with either: (i)* $D_{\mathcal{M}} = 0$ *in which case* $V(\alpha|x, I) = X'\gamma(\alpha|I)$ *as in (2.5); (ii)* $D_{\mathcal{M}} > 0$, *in which case* $V(\alpha|x, I)$ *has* $D_{\mathcal{M}}$ *interactions as in (2.9).*

**Assumption H** *The kernel function* $K(\cdot)$ *with support* $(-1,1)$ *is symmetric, continuously differentiable over the straight line, and strictly positive over* $(-1,1)$. *The positive bandwidth* $h$ *goes to* $0$ *with*

$$\lim_{L \to \infty} \frac{\log L}{Lh^{2(D_{\mathcal{M}}+1)}} = 0.$$

*For the ASQR estimator,* $P(x) = [P_1(x), \ldots, P_K(x)]'$ *where* $P_k(x) = P_{hk}(x)$ *and* $K \asymp h^{-D_{\mathcal{M}}}$. *The retained sieve satisfies the high-level Assumption R stated in Appendix A.*

**Assumption F** *For all* $x$ *in* $\mathcal{X}$ *and* $\alpha$ *in* $[0,1]$, *the function* $\mathcal{F}[\alpha, x, b_{0I}, b_{1I}; I \in \mathcal{I}]$ *is twice differentiable with respect to* $b_{0I}$ *and* $b_{1I}$, $I$ *in* $\mathcal{I}$. *The partial derivatives of order 1 and 2 are continuous with respect to* $\alpha$, $x$, $B_I$ *and* $B_I^{(1)}$, $I$ *in* $\mathcal{I}$.

Assumption A recalls the quantile implications of Bayesian Nash equilibrium bidding under symmetric IPV, see Assumption A-(iii). In Assumption A-(i), the existence of a conditional pdf for the covariate $x_\ell$ is only used for the infinite dimensional quantile regression specification. For a standard quantile regression specification, it is sufficient to assume that the matrix $\mathbb{E}[\mathbb{I}(I_\ell = I) X_\ell X_\ell']$ has an inverse for all $I \in \mathcal{I}$ as recalled in Assumption R-(i) in Appendix A. Note that, as all along this paper, private values and number of bidders can be dependent. A discussion of such dependence in relation with an entry stage preliminary to the auction can be found in Marmer, Shneyerov and Xu (2013a). For Assumption A-(ii), recall that

$$V^{(1)}(\alpha|x, I) = \frac{1}{f(V(\alpha|x, I)|x, I)}, \tag{4.3}$$

where $f(v|x, I)$ is the conditional private value pdf. Hence Assumption A-(ii) amounts to assume that $f(v|x, I)$ is bounded away from 0 and infinity on its support $[V(0|x, I), V(1|x, I)]$ as assumed for instance in Riley and Samuelson (1981), Maskin and Riley (1984) or GPV.

The condition $0 < f(v|x, I) < \infty$ is also used for asymptotic normality of quantile regression estimator, see Koenker (2005). Assumption S combines a standard smoothness assumption with interaction restrictions.

Assumption H restricts the rate at which the bandwidth can go to 0. In the AQR case, it writes $\lim_{L\to\infty} \log L/(Lh^2) = 0$ which is slightly more restrictive than the condition $\lim_{L\to\infty} \log L/(Lh) = 0$ used in nonparametric estimation. This rate restriction is specific to the quantile approach used here. The restriction $K \asymp h^{-D_{\mathcal{M}}}$ and the choice of a sieve satisfying the high-level Assumption R of Appendix A is discussed in the next section.

Assumption F hold for most of the examples of functionals above. A notable exception is the cdf $F(v|x, I)$ in Example 3 when expressed using the rearrangement method of Chernozhukov et al. (2010), which involves an indicator function which is not smooth. However it holds for the smoothed approximation $F_\eta(v|x, I)$ of the cdf, although Assumption F implicitly rules out vanishing bandwidth $\eta$ in Example 3.

### 4.1.2 Choice of a sieve satisfying Assumption H

The last stage of our procedure is the choice of a suitable sieve in (2.10), when a quantile regression specification cannot be used and more flexibility is needed. While the high level Assumption R of Appendix A mentioned in Assumption H describes some key theoretical properties used in the main results, the focus is set here on suitable sieves. The most important requirement is that the sieve has good approximation properties as detailed in Appendix A. Although not strictly necessary, the sieve functions $P_k(\cdot)$ in the private value quantile expansion (2.10) should be localized, i.e. the number of $P_{k'}(\cdot)$ such that $P_k(\cdot)P_{k'}(\cdot)$ do not vanish must be bounded. These two requirements are typically satisfied by sieves building on cardinal spline basis or wavelets as detailed now.

Consider first the spline example of sieves. Assume that $\mathcal{X} = [0,1]^D$ for the sake of brevity. For $m \geq s+2$, set $(t)_+^{m-1} = t^{m-1}$ if $t > 0$ and $(t)_+^{m-1} = 0$ otherwise. The considered spline sieve is based upon the uniformly spaced simple knots $B-$spline function of order $m$

(Schumaker (2007), p.135)

$$q(t)=\sum_{i=0}^{m}\frac{(-1)^{i}\binom{m}{i}(t-i)_{+}^{m-1}}{m!}$$

which has $m-2$ continuous derivatives over the straight line and which support is $[0,m]$. The baseline $B-$spline function $q(\cdot)$ generates the rescaled functions $p_{\kappa h}(\cdot)=p_{\kappa}(\cdot)$

$$p_{\kappa}(t)=\frac{1}{\sqrt{h}}q\left(\frac{t-(\kappa-m)h}{h}\right),\quad \kappa=1,\ldots,\overline{\kappa}$$

where $\overline{\kappa}=\overline{\kappa}_{h}=O(1/h)$ is the largest integer number such that $(\overline{\kappa}-m)h\leq 1\leq\overline{\kappa}h$. Theorem 6.20 in Schumaker (2007) implies that each function $v(\cdot)$ with $s+1$ continuous derivatives can be approximated uniformly over $[0,1]$ with a linear combination of the $p_{\kappa}(\cdot)$'s up to an error $o\left(h^{-(s+1)}\right)$. The $p_{\kappa}(\cdot)$'s are also localized with $\int_{0}^{1}p_{\kappa}^{2}(t)\,dt=O(1)$ uniformly in $\kappa$ and $h$. Similarly, additive quantile functions as in (2.8) can be approximated using the sieve

$$\{p_{\kappa}(x_{1}),\ldots,p_{\kappa}(x_{D}),\kappa=1,\ldots\overline{\kappa}\}.$$

A suitable sieve for additive interactive quantile function of order $D_{\mathcal{M}}$ as in (2.9) is

$$\left\{\prod_{\delta=1}^{D_{\mathcal{M}}}p_{\kappa_{\delta}}(x_{j_{\delta}}),\text{ all }(\kappa_{\delta},j_{\delta})\text{ with }1\leq\kappa_{1},\ldots,\kappa_{D_{\mathcal{M}}}\leq\overline{\kappa},\,1\leq j_{1}<\cdots<j_{\delta}\leq D\right\}.\qquad(4.4)$$

The set (4.4) can be written as a collection $\{P_{k}(x),k=1,\ldots,K\}$ with $K=O\left(h^{-D_{\mathcal{M}}}\right)$ localized functions satisfying $\int_{\mathcal{X}}P_{k}^{2}(x)\,dx=O(1)$ uniformly in $k$ and $h$.

Similar localized sieve can be obtained using wavelets on the interval $[0,1]$, see Härdle, Kerkyacharian, Picard and Tsybakov (1998), Chen (2007) and the references therein, in particular Daubechies (1992). Let $\varphi(\cdot)$ and $\psi(\cdot)$ the father and mother wavelets of order $s+1$, i.e. $\int t^{r}\varphi(t)\,dt=0$ for $r=1,\ldots,s+1$. A wavelet sieve similar to (4.4) is given by the collection of functions

$$\prod_{\delta=1}^{D_{\mathcal{M}}} \frac{1}{2^{-H_0/2}} \varphi\left(\frac{x_{j_\delta} - 2^{-H_0}\kappa_\delta}{2^{H_0}}\right) \text{ and } \prod_{\delta=1}^{D_{\mathcal{M}}} \frac{1}{2^{-H/2}} \psi\left(\frac{x_{j_\delta} - 2^{-H}\kappa_\delta}{2^{H}}\right), \; H_0 \leq H \leq H_1$$

where $H_0$ and $H_1$ are two diverging integer numbers with $2^{-H} \asymp h$, $\kappa_\delta$ and $j_\delta$ as in (4.4).

## 4.2 Private value quantile estimation results

The next sections give our theoretical results for integrated mean squared error and asymptotic distribution of the augmented estimator $\widehat{V}(\cdot|x, I)$. Theorem A.1 in Appendix A also gives uniform consistency rates of similar interest.

### 4.2.1 Integrated mean squared error

Recall $P(x_\ell) = [1, x_\ell']'$ is of the constant dimension $K = D + 1$ in the AQR case. Let $s_1$ be the $1 \times (s+2)$ selection vector $(0, 1, 0, \ldots, 0)$, which is such that $s_1 \otimes \mathrm{Id}_K \widehat{\beta}(\alpha|I) = \widehat{\beta}_1(\alpha|I)$ is the estimator of sieve coefficient derivative $\beta^{(1)}(\alpha)$. Let $\Pi^1(\alpha)$ be the second column of the inverse of $\int \pi(t)\pi(t)' K(t)\,dt$, i.e.,

$$\Pi^1(\alpha) = \left(\int \pi(t)\pi(t)' K(t)\,dt\right)^{-1} s_1'$$

and consider the variance terms

$$\begin{aligned}
v^2(\alpha) &= \Pi^1(\alpha)' \int\int \pi(t_1)\pi(t_2)' \min(t_1, t_2) K(t_1) K(t_2)\,dt_1 dt_2 \Pi^1(\alpha), \\
\Sigma(\alpha|I) &= \frac{\alpha^2 v^2(\alpha)}{(I-1)^2} \mathbb{E}^{-1}\left[\frac{P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] \\
&\quad \times \mathbb{E}\left[P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I)\right] \mathbb{E}^{-1}\left[\frac{P(x_\ell) P(x_\ell)' \mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right], \\
\Sigma_{IL} &= \int_{\mathcal{X}} \int_0^1 P(x)' \Sigma(\alpha|I) P(x)\,d\alpha dx.
\end{aligned}$$

That $v^2(\alpha)$, and then $\Sigma_{IL}$, is strictly positive follows from the proof of Theorem 2 below, see in particular Lemma B.5 in Appendix B. The bias of the estimator will depend upon

$$\begin{aligned}
\mathsf{Bias}(\alpha|I) &= \frac{\alpha}{I-1}s_1\left(\int \pi(t)\,\pi(t)' K(t)\,dt\right)^{-1}\int \frac{t^{s+2}\pi(t)}{(s+2)!}K(t)\,dt \\
&\times \mathbb{E}^{-1}\left[\frac{P(x_\ell)\,P(x_\ell)'\,\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]\mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)\,P(x_\ell)\,\alpha B^{(s+2)}(\alpha|x_\ell, I_\ell)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right], \\
\mathsf{Bias}^2_{IL} &= \int_{\mathcal{X}}\int_0^1 \left(P(x)'\,\mathsf{Bias}(\alpha|I)\right)^2 d\alpha dx.
\end{aligned}$$

**Theorem 2** *Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.5), for which $D_{\mathcal{M}} = 0$, or a sieve quantile regression (2.10) with $D_{\mathcal{M}}$ interactions. Then under Assumptions A, H, S with $s \geq D_{\mathcal{M}}/2$, there exists an approximation $\widehat{v}(\alpha|x, I)$ of $\widehat{V}(\alpha|x, I)$ such that*

$$\begin{aligned}
\mathbb{E}\left[\int_{\mathcal{X}}\int_0^1 \left(\widehat{v}(\alpha|x, I) - V(\alpha|x, I)\right)^2 d\alpha dx\right] &= h^{2(s+1)}\mathsf{Bias}^2_{IL} + \frac{\Sigma_{IL}}{LIh^{D_{\mathcal{M}}+1}} \\
&+ o\left(h^{2(s+1)} + \frac{1}{Lh^{D_{\mathcal{M}}+1}}\right)
\end{aligned}$$

*where $\mathsf{Bias}^2_{IL} = O(1)$, $\Sigma_{IL} = O(1)$ and*

$$\int_{\mathcal{X}}\int_0^1 \left(\widehat{V}(\alpha|x, I) - \widehat{v}(\alpha|x, I)\right)^2 d\alpha dx = o_{\mathbb{P}}\left(\frac{1}{Lh^{D_{\mathcal{M}}+1}}\right). \tag{4.5}$$

The quantile estimator $\widehat{V}(\alpha|x, I)$ is nonlinear and defined in an implicit way, so that attempting a direct computation of its IMSE is difficult. Its approximation $\widehat{v}(\alpha|x, I)$ follows from a Bahadur linearization argument, see Theorem D.1 and (E.1) in Appendices D and E. The rate in equation (4.5) is negligible with respect to the IMSE of $\widehat{v}(\alpha|x, I)$, showing that it is fair to replace $\widehat{V}(\alpha|x, I)$ by $\widehat{v}(\alpha|x, I)$ to picture the IMSE of $\widehat{V}(\alpha|x, I)$.

Note that Theorem 2 holds over the full quantile level range $[0, 1]$. The bias variance decomposition of the IMSE is driven by the estimation of $\alpha B^{(1)}(\alpha|x, I)$ in $V(\alpha|x, I) = B(\alpha|x, I) + \alpha B^{(1)}(\alpha|x, I)/(I-1)$, a function which is $(s+1)$th continuously differentiable which gives the order $h^{s+1}$ for the bias and the order $1/\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}$ for the variance. The bias component due to the estimation of $B(\alpha|x, I)$ is of the negligible order $h^{s+2}$ except perhaps over a small vicinity of 0 where it is $o(h^{s+1})$. The asymptotic variance $\Sigma_{IL}/\left(LIh^{D_{\mathcal{M}}+1}\right)$ order is similar to the asymptotic variance obtained for kernel estimation of a conditional pdf with $D_{\mathcal{M}}$ covariates. Indeed, the bid quantile derivative is homogeneous to a conditional pdf since

$$B^{(1)}(\alpha|x, I) = \frac{1}{g[B(\alpha|x, I)|x, I]},$$

where $g(\cdot|\cdot)$ is the bid conditional pdf. The bid quantile function is homogeneous to a cdf and converges with a faster rate. Note that the asymptotic variance term $\Sigma_{IL}/\left(LIh^{D_{\mathcal{M}}+1}\right)$ depends upon the number of interactions $D_{\mathcal{M}}$ and not the dimension of the covariate $D$. Hence Theorem 2 illustrates the dimension reduction features of the procedure. In particular, the variance term is of order $1/(Lh)$ in the AQR case independently of the dimension of the covariate $D$, which therefore can be large.

Maximizing the leading term of the IMSE yields the optimal bandwidth

$$h_* = \left(\frac{(D_{\mathcal{M}}+1)\,\Sigma_{IL}}{2(s+1)\,\mathsf{Bias}_{IL}^2}\frac{1}{LI}\right)^{\frac{1}{2s+D_{\mathcal{M}}+3}}. \tag{4.6}$$

As in kernel estimation, a pilot bandwidth can be computed using a simple private value quantile regression model to proxy $\Sigma_{IL}$ and $\mathsf{Bias}_{IL}^2$ in a parametric way. The corresponding IMSE rate is

$$L^{\frac{s+1}{2s+D_{\mathcal{M}}+3}}$$

which decreases with the number of interactions $D_{\mathcal{M}}$, but does not depend upon the dimension $D$ of the covariate. In the AQR case with $D_{\mathcal{M}} = 0$, the IMSE rate $L^{\frac{s+1}{2s+3}}$ is, as expected, the optimal rate for estimating the marginal pdf of a real random variable. For $s = 1$, it is equal to $L^{2/5}$ independently of the dimension $D$ of the covariate, which is close of $L^{1/2}$.

Two assumptions limit the use of the optimal bandwidth (4.6). First, Theorem 2 assumes $s \geq D_{\mathcal{M}}/2$ but this condition is only binding for a number of interactions $D_{\mathcal{M}}$ larger than 3 since $s \geq 1$ under Assumption S. Belloni et al. (2017) have a similar restriction for a sieve quantile estimator. In a context where the covariate $D$ replaces $D_{\mathcal{M}}$ but plays a similar role, Aryal et al. (2016) however use a condition $s+1>D$ to study a GMM version of GPV based on a local polynomial estimation of the private value.

### 4.2.2 Central limit theorem

This section states a Central Limit Theorem for $\widehat{V}(\alpha|x,I)$, Theorem 3, which illustrates the good pointwise properties of $\widehat{V}(\alpha|x,I)$ near or at the upper boundary $\alpha=1$. Let $s_1$ be the selection vector defined earlier and

$$\Pi_h^1(\alpha)=\left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi(t)\,\pi(t)'\,K(t)\,dt\right)^{-1}s_1',$$

$$v_h^2(\alpha)=\Pi_h^1(\alpha)'\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi(t_1)\,\pi(t_2)'\min(t_1,t_2)\,K(t_1)\,K(t_2)\,dt_1dt_2\Pi_h^1(\alpha),$$

$$\begin{aligned}\Sigma_h(\alpha|I)=&\frac{\alpha^2v_h^2(\alpha)}{(I-1)^2}\mathbb{E}^{-1}\left[\frac{P(x_\ell)\,P(x_\ell)'\,\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right]\\&\times\mathbb{E}\left[P(x_\ell)\,P(x_\ell)'\,\mathbb{I}(I_\ell=I)\right]\mathbb{E}^{-1}\left[\frac{P(x_\ell)\,P(x_\ell)'\,\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right],\end{aligned}\tag{4.7}$$

$$\begin{aligned}\mathsf{Bias}_h(\alpha|I)=&\frac{\alpha}{I-1}s_1\left(\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\pi(t)\,\pi(t)'\,K(t)\,dt\right)^{-1}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\frac{t^{s+2}\pi(t)}{(s+2)!}K(t)\,dt\\&\times\mathbb{E}^{-1}\left[\frac{P(x_\ell)\,P(x_\ell)'\,\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right]\mathbb{E}\left[\frac{\mathbb{I}(I_\ell=I)\,P(x_\ell)\,\alpha B^{(s+2)}(\alpha|x_\ell,I)}{B^{(1)}(\alpha|x_\ell,I)}\right].\end{aligned}\tag{4.8}$$

**Theorem 3** *Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.5) or a sieve quantile regression (2.10) with $D_{\mathcal{M}}$ interactions. Then under Assumptions A, H, S with $s\geq D_{\mathcal{M}}/2$ and*

$$\frac{\log^2 L}{Lh^{2D_{\mathcal{M}}+1+1\vee D_{\mathcal{M}}}}=o(1),$$

*it holds for $\alpha$ in $(0,1]$ and all $x$ in $\mathcal{X}$ that*

$$\left(\frac{LIh}{P\left(x\right)'\Sigma_h\left(\alpha|I\right)P\left(x\right)}\right)^{1/2}\left(\widehat{V}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)-h^{s+1}P\left(x\right)'\mathsf{Bias}_h(\alpha|I)+o\left(h^{s+1}\right)\right)$$

*converges in distribution to a standard normal. Moreover $P\left(x\right)'\Sigma_h\left(\alpha|I\right)P\left(x\right)\asymp\alpha h^{-D_{\mathcal{M}}}$ and $\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\mathsf{Bias}_h(\alpha|I)\right|=O\left(1\right)$.*

Theorem 3 shows that the asymptotic variance of $\widehat{V}\left(\alpha|x,I\right)$ is of order $\alpha/\left(Lh^{D_{\mathcal{M}}+1}\right)$ for $\alpha>0$. For $\alpha=0$, $\widehat{V}\left(\alpha|x,I\right)=\widehat{B}\left(\alpha|x,I\right)$ has an asymptotic variance of order $1/\left(Lh^{D_{\mathcal{M}}+1}\right)$ and a corresponding CLT using this standardization also holds. For other quantile levels the private value conditional quantile estimator depends upon $\widehat{B}^{(1)}\left(\alpha|x,I\right)$ so that the asymptotic variance of $\widehat{V}\left(\alpha|x,I\right)$ has the larger order $1/\left(Lh^{D_{\mathcal{M}}+1}\right)$ which also holds in Theorem 2. The expression of the asymptotic variance of $\widehat{V}\left(\alpha|x,I\right)$ is quite typical of quantile regression estimators, up to the factor $v_h^2\left(\alpha\right)$ which is due to $\widehat{B}^{(1)}\left(\alpha|x,I\right)$.

It follows from Theorem 3 that the private value conditional quantile estimator is consistent for all quantile levels, including $\alpha=1$. The potential boundary effects only appear through the bias and variance factors $\mathsf{Bias}_h(\alpha|I)$ and $\Sigma_h\left(\alpha|I\right)$. Since the support of the kernel is $[-1,1]$, it holds that

$$\mathsf{Bias}_h(\alpha|I)=\mathsf{Bias}(\alpha|I)\text{ and }\Sigma_h\left(\alpha|I\right)=\Sigma\left(\alpha|I\right)\text{ for all }\alpha\text{ in }[h,1-h]$$

where $\mathsf{Bias}(\alpha|I)$ and $\Sigma\left(\alpha|I\right)$ are defined before Theorem 2, allowing in principle to implement simple pilot bandwidth for quantile level inside $[0,1]$. When $\alpha$ lies in $(0,h]$ or $[1-h,1]$, the bias and variance factors depend upon $h$. It is commonly believed that the variance factor is inflated near the boundaries but there is no clear result for the bias factor, see Fan and Gijbels (1996) and the references therein.

## 4.3 Functional estimation

The plug in estimators of $\theta(x)$ and $\theta$ in (2.13) are

$$\widehat{\theta}(x) = \int_0^1 \mathcal{F}\left[\alpha, x, \widehat{B}(\alpha|x, I), \widehat{B}^{(1)}(\alpha|x, I); I \in \mathcal{I}\right] d\alpha, \quad \widehat{\theta} = \int_{\mathcal{X}} \widehat{\theta}(x)\, dx,$$

with AQR or ASQR $\widehat{B}(\alpha|x, I)$ and $\widehat{B}^{(1)}(\alpha|x, I)$. Alternatively, $\theta$ can be estimated using $\sum_{\ell=1}^{L} \widehat{\theta}(x_\ell)/L$. Let us now introduce the asymptotic variances of $\widehat{\theta}(x)$ and $\widehat{\theta}$. The variances depend upon the matrices

$$\begin{aligned} \mathbf{P}(I) &= \mathbb{E}\left[\mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)\right], \\ \mathbf{P}_0(\alpha|I) &= \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right], \end{aligned}$$

and of the functions, recalling $b_{0I}$ and $b_{1I}$ stand for $B(\alpha|x, I)$ and $B^{(1)}(\alpha|x, I)$ respectively,

$$\begin{aligned} \varphi_{0I}(\alpha, x) &= \frac{\partial \mathcal{F}\left[\alpha, x, B(\alpha|x, I), B^{(1)}(\alpha|x, I); I \in \mathcal{I}\right]}{\partial b_{0I}}, \\ \varphi_{1I}(\alpha, x) &= \frac{\partial \mathcal{F}\left[\alpha, x, B(\alpha|x, I), B^{(1)}(\alpha|x, I); I \in \mathcal{I}\right]}{\partial b_{1I}}. \end{aligned}$$

Let $A$ be a random variable with the uniform distribution over $[0, 1]$ and define

$$\begin{aligned} \sigma_L^2(x|I) &= I \operatorname{Tr}\left\{\operatorname{Var}\left[\left(\int_0^A \left\{\varphi_{0I}(\alpha|x) - \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha}\right\} \mathbf{P}_0(\alpha|I)^{-1} d\alpha\right) \mathbf{P}(I)^{1/2} h^{D_{\mathcal{M}}/2} P(x)\right]\right\}, \\ \sigma_L^2(I) &= I \operatorname{Tr}\left\{\operatorname{Var}\left[\int_0^A \left(\int_{\mathcal{X}} \left\{\varphi_{0I}(\alpha|x) - \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha}\right\} \mathbf{P}_0(\alpha|I)^{-1} \mathbf{P}^{1/2}(I) P(x)\, dx\right) d\alpha\right]\right\}, \\ \sigma_L^2(x) &= \sum_{I \in \mathcal{I}} \sigma_L^2(x|I), \quad \sigma_L^2 = \sum_{I \in \mathcal{I}} \sigma_L^2(I). \end{aligned}$$

The proof of Theorem 4 in Appendix E shows that the asymptotic variances of $\widehat{\theta}(x)$ and $\widehat{\theta}$ are $\sigma_L^2(x)/\left(Lh^{D_{\mathcal{M}}}\right)$ and $\sigma_L^2/L$ respectively provided

$$\varphi_{0I}(\alpha|x) \neq \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha} \tag{4.9}$$

for some $\alpha$, $x$ and $I$ of $[0,1] \times \mathcal{X} \times \mathcal{I}$. Indeed, if $\varphi_{0I}(\alpha|x) = \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha}$ for all $\alpha$ and $I$, $\sigma_L^2(x|I) = 0$ and, if this also holds for all $x$, $\sigma_L^2 = 0$, in which case $\widehat{\theta}(x)$ and $\widehat{\theta}$ can converge to $\theta(x)$ and $\theta$ with "superefficient" rates, faster than $\left(Lh^{D_{\mathcal{M}}}\right)^{1/2}$ and $L^{1/2}$ respectively. In the case of density based functionals, Laurent (1997) similarly obtained asymptotic variance that can vanish. Why it is possible is better understood in our quantile context, through an example of functionals for which (4.9) does not hold.[8] Consider, for some $I_0$ of $\mathcal{I}$,

$$\mathcal{F}_1\left[\alpha, x, B(\alpha|x,I), B^{(1)}(\alpha|x,I); I \in \mathcal{I}\right] = 2B(\alpha|x,I)\, B^{(1)}(\alpha|x,I_0)$$

which gives $(\varphi_{0I_0}(\alpha|x), \varphi_{1I_0}(\alpha|x)) = 2\left(B^{(1)}(\alpha|x,I_0), B(\alpha|x,I)\right)$. Hence $\varphi_{0I}(\alpha|x) = \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha}$ for all $(\alpha, x, I)$, so that (4.9) does not hold and $\sigma_L^2(x) = \sigma_L^2 = 0$. Why $\widehat{\theta}(x)$ and $\widehat{\theta}$ can converge with superefficient rates for these functionals is in fact not surprising observing that they estimate

$$\theta_1(x) = B^2(1|x,I_0) - B^2(0|x,I_0), \quad \theta_1 = \int_{\mathcal{X}} \theta_1(x)\, dx,$$

respectively. Hence, for these examples, the parameters of interest only depend upon extreme quantiles, in which case superefficient estimation is possible, see e.g. Hirano and Porter (2003) and the references therein. A role of the new Condition (4.9) is to exclude such functionals. The next Theorem establishes the asymptotic normality of $\widehat{\theta}(x)$ and $\widehat{\theta}$.

**Theorem 4** *Suppose Assumptions A, F, H, S and R hold with $s \geq D_{\mathcal{M}}/2$. Then $\sigma_L^2(x)$ and $\sigma_L^2$ are bounded away from 0 and infinity if (4.9) holds for some $(\alpha, I)$ in $[0,1] \times \mathcal{I}$ and for some $(\alpha, x, I)$ in $[0,1] \times \mathcal{X} \times \mathcal{I}$ respectively. Moreover*

*i. If $\frac{\log L}{Lh^{2D_{\mathcal{M}}+2+(D_{\mathcal{M}} \vee 1)}} = o(1)$, $\sqrt{Lh^{D_{\mathcal{M}}}}\left(\widehat{\theta}(x) - \theta(x) - \mathsf{bias}_{L,\theta(x)}\right)/\sigma_L(x)$ converges in distribution to a standard normal, where $\mathsf{bias}_{L,\theta(x)}$ is a $o(h^s)$ bias term.*

*ii. If $\frac{\log L}{Lh^{2D_{\mathcal{M}}+1+(D_{\mathcal{M}} \vee 1)}} = o(1)$, $\sqrt{L}\left(\widehat{\theta} - \theta - \mathsf{bias}_{L,\theta}\right)/\sigma_L$ converges in distribution to a standard normal, where $\mathsf{bias}_{L,\theta}$ is a $o(h^s)$ bias term.*

[8] A more systematic study is out of the scope of the present paper, as is the issue of semiparametric efficiency.

The bias term order is given by the estimation of $B^{(1)}(\alpha|x,I)$. When $\mathcal{F}(\cdot)$ depends upon $\alpha B^{(1)}(\alpha|x,I)$ as in all the Examples, the exact order of the bias term is $h^{s+1}$ with

$$\begin{aligned}\mathsf{bias}_{L,\theta(x)} = h^{s+1}(1+o(1))\sum_{i\in\mathcal{I}}\int_0^1 \mathcal{G}_{b_{1I}}\left[\alpha, x, B(\alpha|x,I), \alpha B^{(1)}(\alpha|x,I); I\in\mathcal{I}\right] \\ \times P(x)' \mathsf{Bias}_h(\alpha|x,I)\, d\alpha\end{aligned}$$

and $\mathsf{bias}_{L,\theta} = \int_{\mathcal{X}} \mathsf{bias}_{L,\theta(x)} dx$ where $\mathsf{Bias}_h(\alpha|x,I)$ is as in (4.8) and $\mathcal{G}_{b_{1I}}(\cdot)$ is the partial derivative of $\mathcal{F}(\cdot)$ with respect to $\alpha B^{(1)}(\alpha|x,I)$. $\widehat{\theta}(x)$ or $\widehat{\theta}$ are therefore asymptotically unbiased if $h^{s+1}\sqrt{Lh^{D_{\mathcal{M}}}} = o(1)$ or $h^{s+1}\sqrt{L} = o(1)$ respectively. The items $\mathsf{Bias}_h(\alpha|x,I)$ in the integral expression of $\mathsf{bias}_{L,\theta(x)}$ can be replaced with their limits $\mathsf{Bias}(\alpha|x,I)$ defined before Theorem 2. Theorem 4 applies to our functional Examples as follows.

**Example 1 (cont'd).** Let $\widehat{\theta} = \widehat{\theta}_n/\widehat{\theta}_d$ be the CRRA risk aversion plug in estimator derived from (2.15). Under the bandwidth condition of Theorem 4-(ii), $\widehat{\theta}_n = \theta_n + \mathsf{bias}_{L,\theta_n} + O_{\mathbb{P}}\left(L^{-1/2}\right)$ and $\widehat{\theta}_d = \theta_d + \mathsf{bias}_{L,\theta_d} + O_{\mathbb{P}}\left(L^{-1/2}\right)$. A standard linearization argument then gives that the asymptotic distribution of

$$\sqrt{L}\left(\widehat{\theta} - \frac{\theta_d \mathsf{bias}_{L,\theta_n} - \theta_n \mathsf{bias}_{L,\theta_n}}{\theta_d^2}\right)$$

is the one of

$$\frac{\theta_d\sqrt{L}\left(\widehat{\theta}_n - \theta_n\right) - \theta_n\sqrt{L}\left(\widehat{\theta}_d - \theta_d\right)}{\theta_d^2}$$

which is normal, applying Theorem 4-(ii) with

$$\begin{aligned}&\mathcal{F}\left[\alpha, x, B(\alpha|x,I), B^{(1)}(\alpha|x,I); I\in\mathcal{I}\right] \\ &\quad= \frac{B(\alpha|x,I_1) - B(\alpha|x,I_0)}{\theta_d}\left(\frac{\alpha B^{(1)}(\alpha|x,I_0)}{I_0-1} - \frac{\alpha B^{(1)}(\alpha|x,I_1)}{I_1-1}\right) \\ &\qquad - \frac{\theta_n}{\theta_d^2}\left(\frac{\alpha B^{(1)}(\alpha|x,I_0)}{I_0-1} - \frac{\alpha B^{(1)}(\alpha|x,I_1)}{I_1-1}\right)^2.\end{aligned}$$

The functions $\varphi_{0I}(\alpha|x) - \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha}$ appearing in the asymptotic variances are, for $I = I_1$,

$$
\begin{aligned}
&\varphi_{0I_1}(\alpha|x) - \frac{\partial \varphi_{1I_1}(\alpha|x)}{\partial \alpha} \\
&= \frac{1}{\theta_d}\left(\frac{\alpha B^{(1)}(\alpha|x, I_0)}{I_0 - 1} - \frac{\alpha B^{(1)}(\alpha|x, I_1)}{I_1 - 1}\right) \\
&\quad - \frac{B(\alpha|x, I_1) - B(\alpha|x, I_0) - \alpha\left(B^{(1)}(\alpha|x, I_0) - B^{(1)}(\alpha|x, I_1)\right)}{\theta_d (I_1 - 1)} \\
&\quad + \frac{2\theta_n}{\theta_d^2 (I_1 - 1)}\left(\frac{\alpha B^{(1)}(\alpha|x, I_0)}{I_0 - 1} - \frac{\alpha B^{(1)}(\alpha|x, I_1)}{I_1 - 1}\right) \\
&\quad + \frac{2\theta_n \alpha}{\theta_d^2 (I_1 - 1)}\left(\frac{B^{(1)}(\alpha|x, I_0) + \alpha B^{(2)}(\alpha|x, I_0)}{I_0 - 1} - \frac{B^{(1)}(\alpha|x, I_1) + \alpha B^{(2)}(\alpha|x, I_1)}{I_1 - 1}\right)
\end{aligned}
$$

where $\alpha B^{(2)}(\alpha|x, I)$ is well defined over $[0, 1]$ by (2.3). The case $I = I_0$ is similar. Using these expressions to estimate the asymptotic variance CRRA risk-aversion $\widehat{\theta}$ is difficult due to the second derivative $B^{(2)}(\alpha|x, I)$, which is difficult to estimate. Although not formally studied here, using a bootstrap procedure may be more appropriate.

**Example 2 (cont'd).** Theorem 4-(i) together with Theorem 3 are useful to study the plug in estimator $\widehat{ER}(\alpha_R|x, I)$ derived from (2.17). Theorem 4-(i) gives that the estimator of the integral component $\theta(x; \alpha_R)$ satisfies $\widehat{\theta}(x; \alpha_R) = \theta(x; \alpha_R) + O(h^{s+1}) + O_{\mathbb{P}}\left(1/\sqrt{Lh^{D_{\mathcal{M}}}}\right)$, while Theorem 3 ensures that $\widehat{V}(\alpha|x, I) = V(\alpha|x, I) + O(h^{s+1}) + O_{\mathbb{P}}\left(1/\sqrt{Lh^{D_{\mathcal{M}}+1}}\right)$. As the $O(h^{s+1})$ items correspond to bias terms and the $O_{\mathbb{P}}(\cdot)$ ones are given by the estimation stochastic component, both $\widehat{\theta}(x; \alpha_R)$ and $\widehat{V}(\alpha_R|x, I)$ contribute to the bias of $\widehat{ER}(\alpha_R|x, I)$. The asymptotic distribution of the bias centered $\sqrt{Lh^{D_{\mathcal{M}}+1}}\left(\widehat{ER}(\alpha_R|x, I) - ER(\alpha_R|x, I)\right)$ is the one of $I\alpha_R^{I-1}(1 - \alpha_R)\sqrt{Lh^{D_{\mathcal{M}}+1}}\left(\widehat{V}(\alpha_R|x, I) - V(\alpha_R|x, I)\right)$, which follows from Theorem 3. The uniform consistency Theorem A.1 in Appendix A can be used to study the estimated screening level $\widehat{\alpha}_R(x, I)$ and reserve price $\widehat{V}(\widehat{\alpha}_R(x, I)|x, I)$ obtained by maximizing $\widehat{ER}(\alpha_R|x, I)$.

**Example 3 (cont'd).** Theorem 4-(i) is also useful to study the private value cdf. and pdf, estimator from Example 3, with a fixed bandwidth $\eta$. The proof carries over if $\eta$ goes to 0 with $h = o(\eta)$ and the order of the variance given by Theorem 4-(i) is correct if $\eta$ is of the order of $\eta$. For the cdf estimator $\widehat{F}_{\eta}(v|x,I) = \int_0^1 \mathbb{I}_{\eta}\left[v - \widehat{V}(\alpha|x,I)\right] d\alpha$,

$$\varphi_{0I}(\alpha|x) = -\frac{1}{\eta} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right), \quad \varphi_{1I}(\alpha|x) = \frac{\alpha}{(I-1)\eta} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right),$$
$$\frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha} = \frac{1}{(I-1)\eta} k\left(\frac{v - V(\alpha|x,I)}{\eta}\right) - \frac{\alpha}{(I-1)\eta^2} k^{(1)}\left(\frac{v - V(\alpha|x,I)}{\eta}\right) V^{(1)}(\alpha|x,I).$$

When $\eta$ goes to 0, the dominant part of the variance is, for inner $v$, integrating by parts and setting $V_{x,I} = V(A|x,I)$

$$\begin{aligned}
&\frac{I}{Lh^{D_{\mathcal{M}}}} \operatorname{Tr}\left\{\operatorname{Var}\left[\left(\int_0^A \frac{\partial \varphi_{1I}(\alpha|x)}{\partial \alpha} \mathbf{P}_0(\alpha|I)^{-1} d\alpha\right) \mathbf{P}(I)^{1/2} h^{D_{\mathcal{M}}/2} P(x)\right]\right\} \\
&= \frac{(1+o(1)) I}{Lh^{D_{\mathcal{M}}}} \operatorname{Tr}\left\{\operatorname{Var}\left[\varphi_{1I}(A|x) \frac{\partial \mathbf{P}_0(A|I)^{-1}}{\partial \alpha} \mathbf{P}(I)^{1/2} h^{D_{\mathcal{M}}/2} P(x)\right]\right\} \\
&= \frac{(1+o(1)) I}{(I-1)^2 Lh^{D_{\mathcal{M}}}} \\
&\quad \times \operatorname{Tr}\left\{\operatorname{Var}\left[\frac{F(V_{x,I}|x,I)}{f(V_{x,I}|x,I)} \frac{k\left(\frac{v - V_{x,I}}{\eta}\right)}{\eta} \frac{\partial \mathbf{P}_0(F(V_{x,I}|x,I)|I)^{-1}}{\partial \alpha} \mathbf{P}(I)^{1/2} h^{D_{\mathcal{M}}/2} P(x)\right]\right\} \\
&= \frac{(1+o(1)) I \int k^2(t)\, dt}{(I-1)^2 L\eta h^{D_{\mathcal{M}}}} \left(\frac{F(v|x,I)}{f(v|x,I)}\right)^2 \\
&\quad \times \operatorname{Tr}\left\{\frac{\partial \mathbf{P}_0(F(v|x,I)|I)^{-1}}{\partial \alpha} \mathbf{P}(I)^{1/2} h^{D_{\mathcal{M}}} P(x) P(x)' \mathbf{P}(I)^{1/2} \frac{\partial \mathbf{P}_0(F(v|x,I)|I)^{-1}}{\partial \alpha}\right\}.
\end{aligned}$$

Hence the order of the variance of $\widehat{F}_{\eta}(v|x,I)$ is $1/\left(L\eta h^{D_{\mathcal{M}}}\right)$. Its bias as an estimator of $F(v|x,I)$ has two components: the first is $\mathsf{bias}_{L,F_{\eta}(v|x,I)}$ due to the bias of $\widehat{V}(\alpha|x,I)$ and is of order $O(h^{s+1})$, while the second is $F_{\eta}(v|x,I) - F(v|x,I) = O(\eta^{s+1})$ is $k(\cdot)$ is a kernel of order $s$. It follows that the optimal bandwidths $h$ and $\eta$ must have the same order $L^{-1/(2s+D_{\mathcal{M}}+3)}$ which gives the consistency rate $L^{-(s+1)/(2s+D_{\mathcal{M}}+3)}$. Repeating these steps for the pdf estimator $\widehat{f}_{\eta}(v|x,I)$ gives the same optimal consistency rate $L^{-s/(2s+D_{\mathcal{M}}+3)}$ which, up to a logarithmic term, corresponds to the GPV optimal minimax rate in presence of $D_{\mathcal{M}}$ covariates.

# 5 Simulation experiments

This section reports the results of a simulation experiment for the AQR estimation of the private value quantile function, the expected revenue and optimal reserve price under risk neutrality from first-price auction with $I = 2$. A second simulation experiment considers estimation of risk aversion based on comparison of first-price auctions with $I = 2$ and $I = 3$ as in (2.15) and on comparison with first-price and ascending auctions with $I = 2$. In each case, the considered number of auctions is $L = 100$ and the number of replications is $1,000$.

As the most difficult component to estimate in the private value quantile function is $\alpha B^{(1)}(\alpha|x, I)/(I-1)$, choosing $I = 2$ corresponds to a worst case scenario. By contrast, the simulation experiment in GPV considers $I = 5$ while $I = 3$ or 5 in Marmer and Shneyerov (2012) and Ma, Marmer and Shneyerov (2018). The number of bids in these references range from $1,000$ for GPV to $4,200$ for Marmer and Shneyerov (2012). In a simulation experiment focused on the nonparametric estimation of the utility function of risk averse bidders, Zincenko (2018) considers $I = 2$ with $L = 300$ and $I = 4$ with $L = 150$. Our simulation experiment is therefore more focused on small samples. We also use three covariate while the aforementioned simulation experiments do not consider covariate, with the exception of Zincenko (2018) who increases the number of auctions to $L = 900$ for one or two covariates to cope with the curse of dimensionality.

## 5.1 Model and estimation method

The private value quantile function is given by a quantile regression model with an intercept and three independent covariates with the uniform distribution over $[0, 1]$,

$$V(\alpha|x) = \gamma_0(\alpha) + \gamma_1(\alpha)x_1 + \gamma_2(\alpha)x_2 + \gamma_3(\alpha)x_3$$

with

$$\gamma_0(\alpha) = 1 + 0.5\exp(5(\alpha - 1)), \quad \gamma_1(\alpha) = 1,$$
$$\gamma_2(\alpha) = 0.5(1 - \exp(-5\alpha)), \quad \gamma_3(\alpha) = 0.8 + 0.15((2\pi + 1)\alpha + \cos(2\pi\alpha)).$$

The coefficient $\gamma_0(\cdot)$ is flat near 0 and fastly increases near 1, as observed in the application displayed in the next section, while $\gamma_2(\cdot)$ fastly increases near 0 and is flat after. The derivative of $\gamma_3(\cdot)$ has some oscillating patterns.

The expected revenue $ER(\alpha)$ is computed from (2.17) setting the intercept, $x_1$ and $x_3$ to 0 and taking $x_2 = 0.8$. This choice gives a unique optimal reserve price achieved for $\alpha = .3$, which is not too close to the boundaries so that the expected revenue function has a substantial concave shape which is suppose to make estimation more difficult.

## 5.2 Private value and expected revenue

The private value quantile regression is estimated from a sample of 100 first-price auctions with two bids over the estimation grid $\alpha = 0, 0.01, \ldots, 0.99, 1$ with an augmented quantile regression estimator $\widehat{V}(\alpha|x)$ of order 2 and kernel $K(t) = 6t(1-t)\mathbb{I}(t \in [0,1])$. The expected revenue estimator $\widehat{ER}(\alpha)$ plugs $0.8\widehat{\gamma}_2(\alpha)$ into (2.17) using Riemann sums to compute integrals. The optimal screening level $\widehat{\alpha}_*$ maximizes $\widehat{ER}(\alpha)$ over the grid and is used to compute the estimated optimal reserve price $\widehat{R}_* = .8\widehat{\gamma}_2(\widehat{\alpha}_*)$ and the estimated optimal revenue $\widehat{ER}_* = \widehat{ER}(\widehat{\alpha}_*)$.

Table 1 summarizes the simulation results for the estimation of the private value quantile function, the expected revenue and the optimal reserve price. The Bias and Square Root Integrated Mean Squared Error (RIMSE) lines for $\widehat{V}(\cdot|\cdot)$ gives the simulation counterparts of, respectively

$$\left(\frac{1}{4}\sum_{j=0}^{3}\int_0^1 \left(\mathbb{E}\left[\widehat{\gamma}_j(\alpha)\right] - \gamma_j(\alpha)\right)^2 d\alpha\right)^{1/2} \text{ and } \left(\frac{1}{4}\sum_{j=0}^{3}\int_0^1 \mathbb{E}\left[\left(\widehat{\gamma}_j(\alpha) - \gamma_j(\alpha)\right)^2\right] d\alpha\right)^{1/2}.$$

The Bias and RIMSE for the expected revenue are computed similarly. Table 1 also gives the Bias and Mean Squared Error (MSE) of the optimal reserve price estimator. All these quantities are computed for bandwidths $.2, .3, \ldots, .9$.

| | $h$ | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 |
|---|---|---|---|---|---|---|---|---|---|
| $\widehat{V}(\cdot\|\cdot)$ | Bias | .131 | .141 | .143 | .145 | .150 | .159 | .166 | .176 |
| | RIMSE | .433 | .386 | .355 | .332 | .322 | .309 | .303 | .305 |
| $\widehat{ER}(\cdot)$ | Bias | .036 | .044 | .049 | .050 | .051 | .049 | .047 | .045 |
| | RIMSE | .109 | .104 | .102 | .100 | .099 | .098 | .097 | .096 |
| $\widehat{R}_*$ | Bias | -.036 | -.031 | -.014 | -.002 | .009 | .022 | .037 | .043 |
| | RMSE | .129 | .099 | .075 | .067 | .062 | .064 | .066 | .066 |

Table 1: Private value quantile function, expected revenue, and optimal reserve price

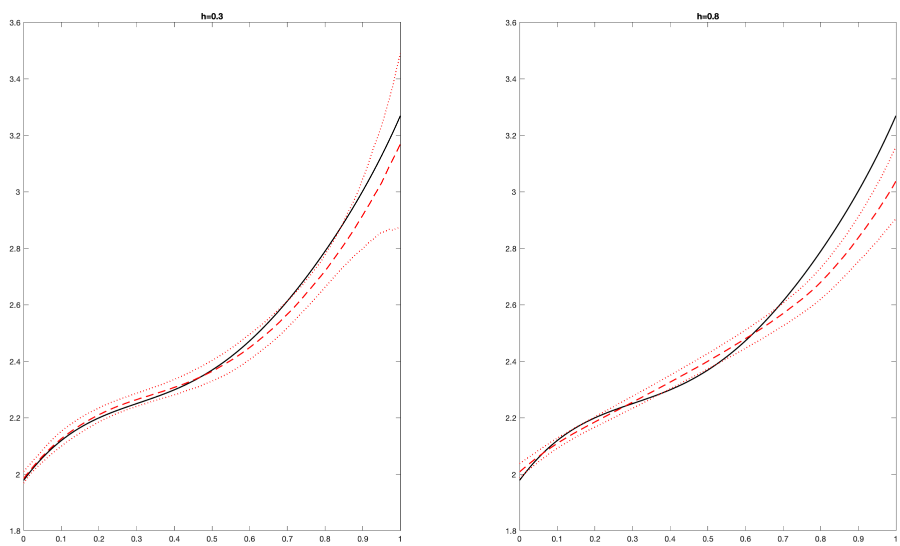


Figure 2: Private value quantile estimation for $h = 0.3$ (left) and $h = 0.8$ (right) for average covariate. True $V(\alpha|x) = \gamma_0(\alpha) + (\gamma_1(\alpha) + \gamma_2(\alpha) + \gamma_3(\alpha))/2$ in black. Dashed red line: average estimation. Dotted red line: pointwise $2.5\% - 97.5\%$ quantiles of $\widehat{V}(\alpha|x)$ across $1,000$ simulations.

Estimation of the private value slope coefficients seems much more sensitive to the bandwidth parameter than the expected revenue or optimal reserve price. It has also a much higher RIMSE. The bandwidth behavior of $\widehat{V}(\alpha|x)$ is illustrated in Figure 2, which consid-

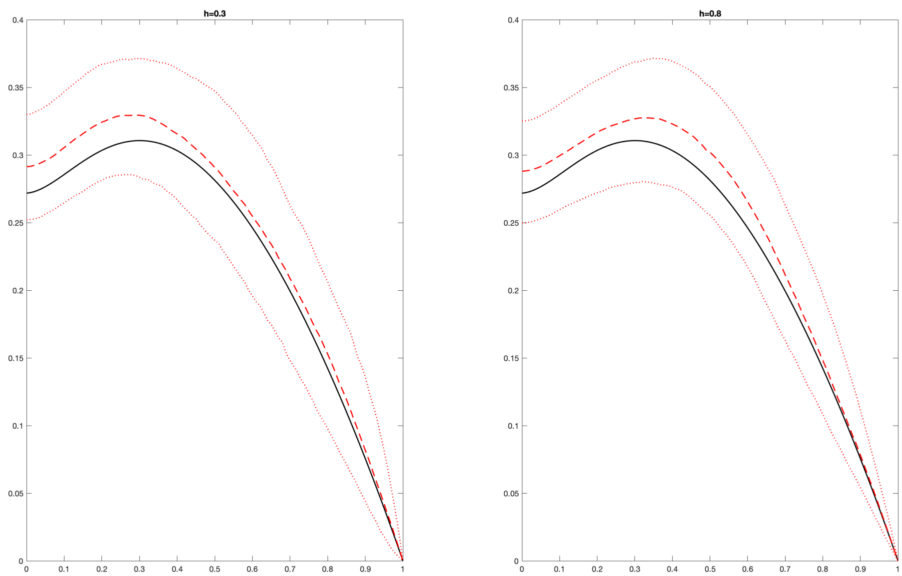


Figure 3: Expected revenue estimation for $h = 0.3$ (left) and $h = 0.8$ (right). True $ER\,(\alpha|x)$ in black. Dashed red line: average estimation. Dotted red line: pointwise $2.5\% - 97.5\%$ quantiles of $\widehat{ER}\,(\alpha|x)$ across $1,000$ simulations.

ers the small bandwidth $h = 0.3$ and the larger $h = 0.8$. As expected from Theorem 3, the variance of $\widehat{V}\,(\alpha|x)$ increases with $\alpha$ and decreases with $h$, while the bias increases with $\alpha$ but decreases with $h$. Figure 2 also suggests that choosing a large bandwidth as recommended by Table 1 may lead to important bias issues, including underestimating the private value quantile function for high $\alpha$.

This contrasts with estimation of the expected revenue and optimal reserve price, which seems mostly unaffected by the bandwidth. This is because the expected revenue depends upon $(1-\alpha)\,V\,(\alpha|x)$: multiplying the private value quantile function by $(1-\alpha)$ mitigates larger bias and variance near the boundary $\alpha = 1$, see also Figure 3. For the considered experiment, the true expected revenue is always in the 95% band of Figure 3 while the true private quantile function is out for large $\alpha$ when $h = 0.8$.

### 5.3 CRRA risk aversion

Two risk aversion estimators are considered. The first estimator $\widehat{\theta}_{fp}$ is based upon (2.15) and uses two independent samples of size $L = 100$ with 2 and 3 bidders from the model above, which corresponds to a CRRA utility function $x^{\theta}$ with $\theta = 1$.[9] Integrals with respect to $\alpha$ are computed using Riemann sums whereas integrals with respect to $x$ are replaced with sample means over the two auction samples. The second estimator $\widehat{\theta}_{asc}$ is based upon (2.16) and uses an additional sample of size $L = 100$ of ascending auctions with two bidders. In this case, it is possible to consider various values of $\theta$ and the simulation experiment considers the values 0.2, 0.6 and 1. Indeed, if $B(\alpha|x)$ is the first-price auction quantile bid function with $I = 2$, the observed bids drawn from $B(\alpha|x)$ are rationalized by a CRRA utility function $x^{\theta}$ if the private value quantile function is set to

$$V_{\theta}(\alpha|x) = B(\alpha|x, 2) + \theta\alpha B^{(1)}(\alpha|x, 2)$$

provided $V_{\theta}^{(1)}(\cdot|x) > 0$ for all $x$ as seen from Campo et al. (2011) and (2.14) here. As $V_{\theta}^{(1)}(\cdot|\cdot) > 0$ holds in our case, we use $V_{\theta}(\alpha|x)$ to generate two ascending bids for each auction. Following Gimenes (2017), $V_{\theta}(\alpha|x)$ can be estimated from winning bids in these ascending auction using AQR for quantile level $2\alpha - \alpha^2$ instead of $\alpha$.

The performance of the two estimators are summarized in the next Table. Table 2 shows

| | $\theta$ | $h$ | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\widehat{\theta}_{fp}$ | 1 | Bias | -.795 | -.564 | -.412 | -.288 | -.178 | .-.080 | .003 | .053 |
| | | RMSE | .891 | .681 | .545 | .471 | .404 | .380 | .393 | .436 |
| $\widehat{\theta}_{asc}$ | 1 | Bias | -.016 | -.019 | -.037 | -.061 | -.085 | -.100 | -.109 | -.111 |
| | | RMSE | .240 | .247 | .248 | .254 | .260 | .267 | .276 | .282 |
| | .6 | Bias | .028 | .023 | .009 | -.008 | -.025 | -.035 | -.040 | -.042 |
| | | RMSE | .172 | .176 | .174 | .175 | .175 | .179 | .184 | .188 |
| | .2 | Bias | .088 | .083 | .075 | .066 | .058 | .053 | .052 | .053 |
| | | RMSE | .135 | .133 | .126 | .122 | .117 | .116 | .116 | .118 |

Table 2: Risk aversion estimation

[9]The optimal bid functions can be computed explicitly under the risk neutrality case $\theta = 1$. Considering other values of $\theta$ would request to use numerical computations of the bid functions.

that $\widehat{\theta}_{asc}$, which combines first-price and ascending auctions as in Lu and Perrigne (2008), dominates $\widehat{\theta}_{fp}$ in this experiment. While the RMSE and bias of $\widehat{\theta}_{asc}$ do not seem sensitive to $h$, this is not the case for $\widehat{\theta}_{fp}$ which has a high downward bias, and then RMSE, for small $h$. Further investigations suggest this is due to an unbalanced variable issue, the difference $\widehat{B}(\alpha|x,3) - \widehat{B}(\alpha|x,2)$ being very smooth while $\alpha\left(\widehat{B}^{(1)}(\alpha|x,3)/2 - \widehat{B}^{(1)}(\alpha|x,2)\right)$ is more erratic, especially when $\alpha$ is close to 1. This issue is addressed in the application by restricting $\alpha$ to $[0,.8]$ for risk aversion estimation.

# 6 Empirical application

This section illustrates empirically the methodology using data from ascending timber auctions run by the US Forest Service (USFS). Timber auctions data have been used in several empirical studies (see Athey and Levin (2001), Athey, Levin and Seira (2011) Li and Zheng (2012), Aradillas-Lopez, Gandhi and Quint (2013) among others). Some other works have investigated risk-aversion on timber auctions (e.g., Lu ad Perrigne (2008), Athey and Levin (2001), Campo et al. (2011)). The data set used here is from Lu and Perrigne (2008) and Campo et al. (2011), and aggregates auctions from the states covering the western half of the United States (regions 1–6 as labeled by the USFS) occurred in 1979. It contains bids and a set of variables characterizing each timber tract, including the estimated volume of the timber measured in thousands of board feet (mbf) and its estimated appraisal value given in dollars per unit of volume. We consider the 107 first-price auctions with two bidders, the first-auctions with three bidders ($L = 108$) and ascending auctions with two bidders ($L = 241$). The considered covariates are the appraisal value and the timber volume taken in log. The rest of the application uses a quantile regression model for the private value, which is estimated via AQR of order 2 and kernel $K(t) = 6t(1-t)\mathbb{I}(t \in [0,1])$, for bandwidths $h$ in $\{.2, .3, \ldots, .9\}$. Confidence intervals are computed using pairwise bootstrap.

**Bid quantile functions.** Table 3 gives the coefficients of a regression on these variables. The dependent variables are the bids for the first-price auctions while the winning bid is used for the ascending auction. The appraisal value coefficient is close to 1 in all auctions, but

| Auctions | | Intercept | Volume | Appraisal value | $R^2$ |
|---|---|---|---|---|---|
| First-price | $I = 2$ | $-1.06$ (6.67) | 4.07 (1.12) | 1.01 (0.04) | 0.77 |
| | $I = 3$ | $-20.79$ (9.55) | 7.10 (1.34) | 1.15 (0.06) | 0.70 |
| Ascending | $I = 2$ | 2.76 (15.05) | 3.76 (1.85) | 1.12 (0.06) | 0.67 |

Table 3: Auction bid regressions

is found significantly distinct at the 5% level when comparing the first-price auction with $I = 2$ with the one with $I = 3$ and the ascending auction. Similarly the volume coefficient of the first-price auction with $I = 2$ differs from the one with $I = 3$ at the 10% level, and also at the 5% level when using an unilateral test. These findings are consistent with a quantile regression specification with non constant coefficients for these two variables. The intercept coefficients of the first-price auctions with $I = 2$ and $I = 3$ are not statistically distinct at the 5% level. This is not compatible with the homogenized bid regression model $V = \gamma_0 + X'\gamma_1 + v$ with $v$ independent of $X$: for this model, the volume and appraisal value coefficients obtained from a bid regression should not depend upon $I$ under entry exogeneity, as discussed in Section 2.2.

Figure 4 sums up the quantile regression analysis of the first-price auction bids with $I = 2$. The difference of the AQR volume slope and regression coefficient is consistently outside the pointwise 90% bootstrap confidence interval. This finding holds for all bandwidths. The case of the appraisal value is more difficult. The differences between the estimated regression coefficient and the AQR lies outside the confidence bands between $\alpha = 90\%$ and $\alpha = 1$ due to a strong increase of the AQR. But this holds for the bandwidths $h = .2$ and $h = .3$ and not for larger $h$. Figure 4 also reports standard quantile regression, which exhibits a similar pattern. This suggests a potential AQR bias issue for $h > .3$.

The intercept slope does not look significant except may be for large $\alpha$ as suggested by the standard QR estimation. Therefore, the intercept will be kept constant and set to its estimated value from Table 3 in the rest of the application. Comparison of the augmented

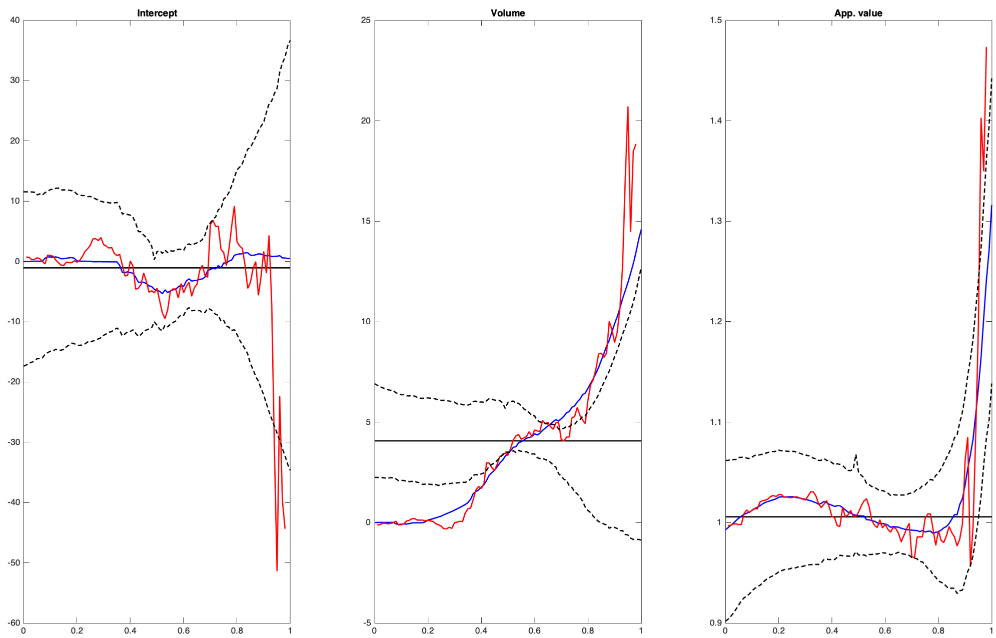


Figure 4: Two bidders first-price auction bid quantile slope coefficients: Intercept (left), volume (center) and appraisal value (right). AQR with $h = .3$ (blue), standard QR (red) and OLS regression (black), and pointwise 90% confidence intervals for the AQR-regression difference (black dashed line) centered at the regression coefficients. A regression or AQR estimated slope coefficients outside the confidence band indicates a potential misspecification of the homogenized bid regression model.

and standard quantile regression estimation also shows that the former produces much more regular slope coefficients.

**Risk aversion.** The two risk aversion estimators look insensitive to the bandwidth, producing a risk aversion estimation around .85 for $\widehat{\theta}_{fp}$ and .7 for $\widehat{\theta}_{asc}$. The bootstrap median of $\widehat{\theta}_{fp}$ reported in Table 4 suggests that the distribution of $\widehat{\theta}_{fp}$ is asymmetric, with a median around .75 slightly above the one of $\widehat{\theta}_{asc}$. This risk aversion estimates are similar to the ones obtained with Lu and Perrigne (2008) and Campo et al. (2011). The bootstrap 90% confidence intervals in Table 4 suggests a much higher dispersion that the ones reported by these authors from asymptotic variance estimations. In particular, it is not possible to reject risk neutrality.

| | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 |
|---|---|---|---|---|---|---|---|---|
| $\widehat{\theta}_{fp}$ (50%) | .92(.72) | .82(.69) | .83(.69) | .84(.71) | .87(.73) | .87(.75) | .86(.77) | .86(.78) |
| [5%, 95%] | [.2, 1.6] | [.2, 1.4] | [.2, 1.4] | [.2, 1.4] | [.2, 1.4] | [.2, 1.4] | [.2, 1.4] | [.2, .1.4] |
| $\widehat{\theta}_{asc}$ (50%) | .75(.72) | .68(.66) | .67(.64) | .67(.64) | .67(.64) | .67(.65) | .66(.65) | .66(.65) |
| [5%, 95%] | [.3, 1.3] | [.3, 1.2] | [.3, 1.3] | [.3, 1.3] | [.3, 1.2] | [.3, 1.2] | [.3, 1.2] | [.3, 1.2] |

Table 4: Risk aversion estimations with $5\%, 50\%$ and $95\%$ bootstrap quantiles, $h = .2, ..., .9$.

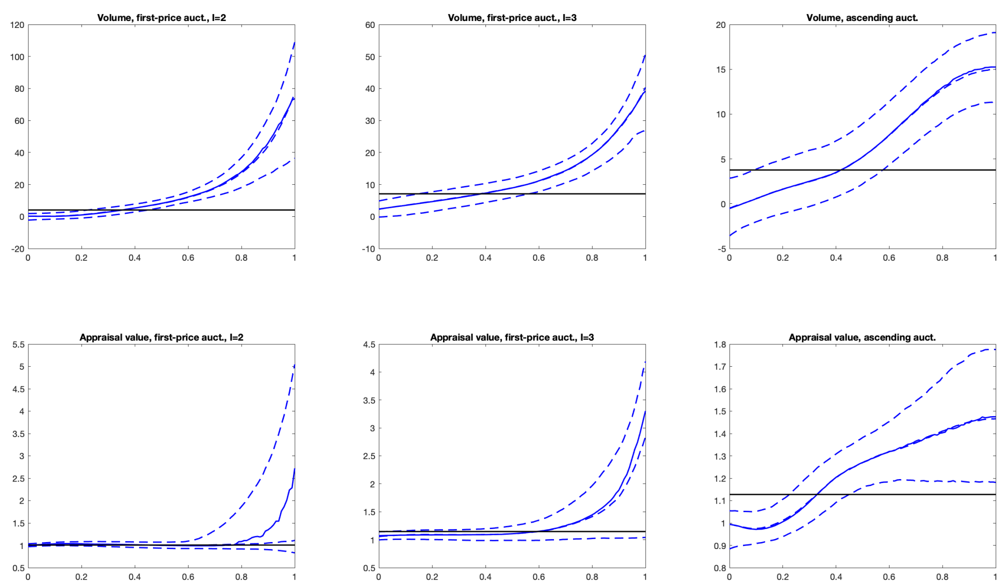


Figure 5: Volume (top) and appraisal value (bottom) estimated private value slope function for first-price auctions with two bidders (left), three bidders (center) and ascending auctions (right), for $h = .3$. AQR estimation (full line), regression (full straight line) and $5\%, 50\%, 95\%$ bootstrapped quantile (dashed line).

**Private value quantile function and expected revenue.** This section reports estimation results under risk neutrality for first-price auctions with two and three bidders and ascending auctions. Figure 5 gives the private value slope function of the volume and appraisal variables. The volume slope functions differ of the corresponding OLS coefficients for all auctions and all the considered bandwidths. Its shape however varies across auctions:

while convex and in the $[20, 100]$ range for high $\alpha$ in the first-price case, it is in the $[8, 15]$ range and more oscillating for ascending auctions. This suggests that the private value distribution is not independent of the auction mechanism.

The appraisal value slope seems statistically different from its OLS counterpart for ascending auctions and may be for first-price auctions with three bidders. For all auctions, the estimated appraisal value slopes start at 1 for $\alpha$ near 0, suggesting that low type bidder valuations of timber lots are very close to the appraisal value. This contrasts with high type bidders with higher $\alpha$, which markup can be very high, in a significant way for the case of ascending auction. This illustrates again the important difference between low type and high type bidders.

A possible discrepancy between first-price and ascending auctions with two bidders also appears in the expected revenue computed for median values of the two explanatory variables, see Figure 6.

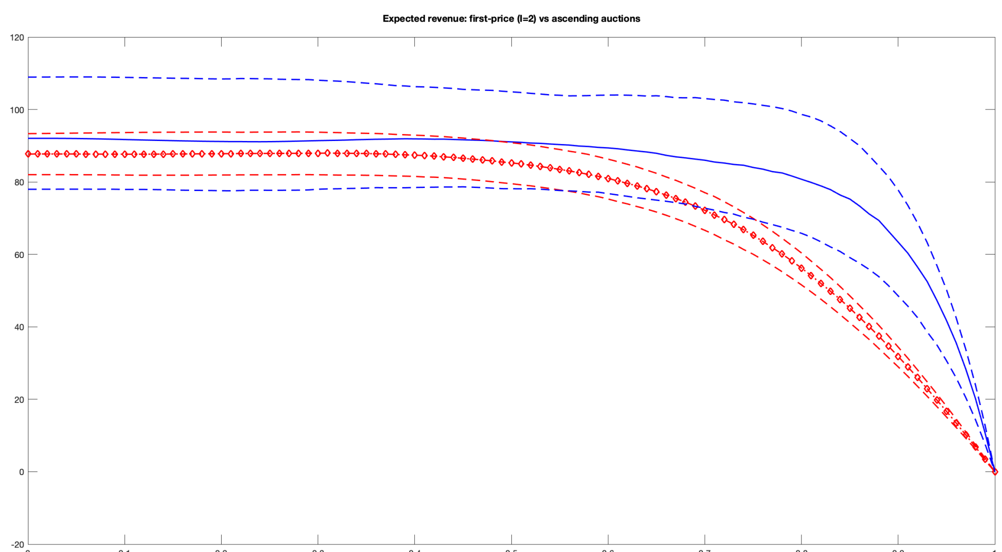


Figure 6: Estimated expected revenue for first-price (full line) and ascending (diamond) auctions with two bidders ($h = .4$). Volume and appraisal value set to median of the first-price auctions. $5\% - 95\%$ bootstrap quantiles in dashed lines.

The ascending auction expected revenue is always below the first-price one. This seems statistically significant for high quantile levels. This may not be relevant for the seller as the optimal revenue is achieved for a wide range $[0, .5]$ of quantile levels over which the two expected revenue curves seem flat. This feature, which appears for all the considered bandwidths, suggests again that the private value quantile function of bidders participating to first-price auctions is higher than the one for ascending auctions. Note also that the bootstrap confidence bands for first-price auction are larger than for ascending one, as for all the estimations reported in this application.

# 7 Conclusion

This paper has presented a quantile regression modeling strategy for first-price auction with risk neutral bidders under the independent value paradigm. For a conditional private value quantile function given by a quantile regression, the conditional bid quantile function is also a quantile regression. Detecting the quantile regression slope which are not constant can be done looking for the corresponding bid quantile regression slope, or with less rigor to the variation of the corresponding homogenized bid regression coefficient with respect to the number of bidders, which is also a consistent estimator of the constant private value slope. Non constant private value slope functions can be recovered from their bid counterparts and their derivative with respect to quantile level. The latter can be estimated using the augmented quantile regression proposed in this paper, which applies local polynomial to estimate jointly the bid quantile slope and its derivatives. This approach is found to work well both in simulations, and in a timber auction application where a strong low type/high type bidder heterogeneity is detected. This can be interpreted as caused by heterogeneous bidder abilities to transform the auctioned timber lots into more valuable goods. An empirical finding is that the seller expected revenue in a median auction is higher in first-price than in ascending auctions. The estimated expected revenue curves look flat for reserve prices below a quite large threshold, including optimal ones. This suggests that the choice of a reserve price may not be important, at least for the median auction considered in the application.

A new local polynomial estimation procedure for bid quantile regression and its quantile level derivatives is proposed to implement this strategy. It is based on a smoothed objective function which produces smooth estimations as illustrated in the simulations and the empirical applications. The auction modeling strategy also applies for unspecified quantile functions thanks to linear sieve methods. This also allows to consider flexible and parsimonious specification such as additive quantile function. The proposed private value quantile estimator converges with nonparametric rates, mimicking the fast optimal ones achieved in the absence of covariate for a quantile regression, or for univariate covariate for an additive quantile specification. Various functionals of private value quantile functions are considered, such as the expected revenue, the private value conditional cdf and pdf, or risk-aversion for bidders with a common CRRA utility function.

Many work remain to be done. The asymptotic distribution derived for the proposed estimators often have a complicated variance, so it may be wiser to use bootstrap inference as in the empirical application. The risk aversion exhibits a quite large variance, suggesting that a better understanding of efficiency issues is needed. Various extensions can also be considered. The quantile approach can be extended to exchangeable affiliated values as considered in Hubbard, Li and Paarsch (2012). The quantile regression with unobserved variables estimation method of Wei and Carroll (2009) can be used to tackle unobserved heterogeneity as in Krasnokutskaya (2011). The quantile identification and estimation strategy can be modified to tackle with endogenous entry, such as reserve price as in Guerre, Perrigne and Vuong (2000) or entry costs as considered by Marmer, Shneyerov and Xu (2013a) or Gentry and Li (2014).

# References

[1] Andrews, D.W.K & Y.J. Whang (1990). Additive interactive regression models: circumventing the curse of dimensionality. *Econometric Theory* **6**, 466–479.

[2] Athey, S. & J. Levin (2001). Information and competition in U.S. Forest Service timber auctions. *Journal of Political Economy* **109**, 375–417.

[3] Athey, S., J. Levin & E. Seira (2011). Comparing open and sealed bid auctions: evidence from timber auctions. *The Quarterly Journal of Economics* **126**, 207–257.

[4] Aradillas-Lopez, A., A. Gandhi & D. Quint (2013). Identification and inference in ascending auctions with correlated private values. *Econometrica* **81,** 489–534.

[5] Aryal, G., M.F. Gabrielli & Q. Vuong (2016). Semiparametric estimation of first-price auction models. CONICET and Universidad Nacional de Cuyo, University of Virginia.

[6] Bassett, G. & R. Koenker (1982). An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association,* **77**, 407–415.

[7] Belloni, A., V. Chernozhukov, D. Chetverikov & I. Fernandes-Vál (2017). Conditional quantile processes based on series or many regressors. arXiv:1105.6154v3.

[8] Campo, S., E. Guerre, I. Perrigne & Q. Vuong (2011). Semiparametric estimation of first-price auctions with risk-averse bidders. *Review of Economic Studies,* **78**, 112–147.

[9] Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. Chap. 76 in *Handbook of Econometrics,* Vol. **6B**. Elsevier.

[10] Chernozhukov, V., I. Fernándes-Val & A. Galichon (2010). Quantile and probability curves without crossing. *Econometrica* **78**, 1093-1125.

[11] Daubechies, I. (1992). *Ten lectures on wavelets.* SIAM.

[12] Dette, H. & S. Volgushev (2008). Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society: Series B* **70**, 609–627.

[13] Enache, A. & J.P. Florens (2015). A quantile approach for the estimation of first-price private value auction. Working Paper, Paris School of Economics.

[14] Fan, J. & I. Gijbels (1996). *Local polynomial modeling and its applications.* Chapman and Hall/CRC.

[15] Gentry, M. & T. Li (2014). Identification in auctions with selective entry. *Econometrica* **82**, 315–344.

[16] Gimenes, N. (2017). Econometrics of ascending auction by quantile regression. *Review of Economics and Statistics* **99**, 944–953.

[17] Guerre, E., I. Perrigne & Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica* **68**, 525–574.

[18] Guerre, E., I. Perrigne & Q. Vuong (2009). Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica* **77**, 1193–1227.

[19] Guerre, E. & C. Sabbah (2012). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory* **28**, 87–129.

[20] Härdle, W., G. Kerkyacharian, D. Picard & A. Tsybakov (1998). *Wavelets, approximation and statistical applications.* Springer.

[21] Haile, P.A., H. Hong & M. Shum (2003). Nonparametric tests for common values in first-price sealed-bid auctions. Cowles Foundation discussion paper.

[22] Hickman, B.R. & T.P. Hubbard (2015). Replacing sample trimming with boundary correction in nonparametric estimation of first-price auctions. *Journal of Applied Econometrics*, **30**, 736-762.

[23] Hirano, K. & J.R. Porter (2003). Asymptotic efficiency in parameter structural models with parameter-dependent support. *Econometrica*, **71**, 1307–1338.

[24] Horowitz, J.L. & S. Lee (2005). Nonparametric estimation of an additive quantile regression model. *Journal of the American Statistical Association* **100**, 1238–1249.

[25] Hubbard, T.P., T. Li & H.J. Paarsch (2012). Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. *Journal of Econometrics*, **168**, 4–16.

[26] Koenker, R. (2005). *Quantile regression.* Cambridge University Press.

[27] Koenker, R. & G. Bassett (1978). Regression quantiles. *Econometrica,* **46**, 33–50.

[28] Krasnokutskaya, E. (2011). Identification and estimation of auctions models with unobserved heterogeneity. *Review of Economic Studies,* **78**, 293–327.

[29] Laurent, B. (1997). Estimation of integral functionals of a density and its derivatives. *Bernoulli,* **3**, 181–211

[30] Li, T., I. Perrigne & Q. Vuong (2003). Semiparametric estimation of the optimal reserve price in first-price Auctions, *Journal of Business & Economic Statistics* **21**, 53–64.

[31] Li, T. & X. Zheng (2012). Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics* **168**, 29–46.

[32] Liu, N. & Y. Luo (2017). A nonparametric test of exogenous participation in first-price auctions. *International Economic Review* **58**, 857–887

[33] Liu, N. & Q. Vuong (2018). Nonparametric test of monotonicity of bidding strategy in first price auctions. Working paper.

[34] Lu, J. & I. Perrigne (2008). Estimating risk aversion from ascending and sealed-bid auctions: the case of timber auction data. *Journal of Applied Econometrics* **23**, 871–896.

[35] Luo, Y. & Y. Wan (2018). Integrated-Quantile-Based Estimation for First-Price Auction Models. *Journal of Business & Economic Statistics* **36**, 173-180.

[36] Ma, J., Marmer, V., & A. Shneyerov (2018). Inference for first-price auctions with Guerre, Perrigne and Vuong's estimator. Working paper,

[37] Marmer, V., & A. Shneyerov (2012). Quantile-based nonparametric inference for first-price auctions. *Journal of Econometrics* **167**, 345–357.

[38] Marmer, V., A. Shneyerov & P. Xu (2013a). What model for entry in first-price auctions? A nonparametric approach. *The Journal of Econometrics*, **176**, 46–58.

[39] Marmer, V., A. Shneyerov & P. Xu (2013b). What model for entry in first-price auctions? A nonparametric approach. Supplementary material. Website of *The Journal of Econometrics.*

[40] Maskin, E. & J.G. Riley (1984). Optimal auctions with risk averse buyers. *Econometrica* **52,** 1473–1518.

[41] Menzel, K. & P. Morganti (2013). Large sample properties for estimators based on the order statistics approach in auctions. *Quantitative Economics,* **4,** 329–375.

[42] Milgrom, P.R. (2001). *Putting auction theory to work.* Cambridge University Press.

[43] Milgrom, P.R. & R.J. Weber (1982). A theory of auctions and competitive bidding. *Econometrica*, **50**, 1089–1122.

[44] Paarsch, H.J. & H. Hong (2006). *An introduction to the structural econometrics of auction data.* MIT Press.

[45] Rezende, L. (2008). Econometrics of auctions by least squares. *Journal of Applied Econometrics* **23**, 925–948.

[46] Schumaker, L.L. (2007). *Spline functions: basic theory.* Cambridge University Press.

[47] Wei, Y. & R.J. Carroll (2009). Quantile regression with measurement error. *Journal of the American Statistical Association* **104**, 1129–1143.

[48] Zincenko, F. (2018). Nonparametric estimation of first-price auctions with risk-averse bidders. *Journal of Econometrics* **205**, 303–335.

# Online Appendix A: Sieve assumption and uniform consistency results

## A.1 High-level sieve assumption

Section 4.1.2 suggests to use spline or wavelet but our results hold for more general sieve choices satisfying the high level Assumption R. The first key condition is the following approximation property.

**Approximation property S**. *For each function $V(\alpha; x)$ with $D_{\mathcal{M}}$ interactions as in (2.9), $(s+1)$th continuously differentiable over $[0,1] \times X$, there exists some coefficients $\gamma_k(\cdot) = \gamma_{kK}(\cdot)$, $(s+1)$th continuously differentiable over $[0,1]$ with equicontinuous $\gamma_{kK}^{(s+1)}(\cdot)$, such that*

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| V(\alpha; x) - \sum_{k=1}^{K} \gamma_k(\alpha) P_k(x) \right| = o\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right), \tag{A.1.1}$$

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| \frac{\partial^p V(\alpha; x)}{\partial \alpha^p} - \sum_{k=1}^{K} \gamma_k^{(p)}(\alpha) P_k(x) \right| = o(1), \quad p = 1, \ldots, s+1. \tag{A.1.2}$$

Note that $K \asymp 1/h$ under Assumption H. Chen (2007) gives a $O\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right)$ rate for standard sieve methods and functions with $s+1$ bounded derivatives, which is comparable to rate in (A.1.1). The rate $o\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right)$ holds for functions with continuous derivatives of order $s+1$ for multivariate B splines (Schumaker, 2007) of order $s+1$ as in (4.4), or multivariate wavelets generated by a father wavelet $p(\cdot)$ function of order $s+1$, see Härdle et al. (1998), Chen (2007) and the references therein, in particular Daubechies (1992). These two sieve also satisfy (A.1.2) as the corresponding coefficients $\gamma_k(\cdot)$ can be written as $\int_{\mathcal{X}} \lambda_k(x) V(\alpha; x)\, dx$ for well chosen $\lambda_k(\cdot) = \lambda_{kK}(\cdot)$ satisfying $\sup_K \int_{\mathcal{X}} |\lambda_k(x)|\, dx < \infty$. The high-level sieve assumption considered in our results is as follows.

**Assumption R** *The sieve satisfies the Approximation property S. In the AQR case the matrices $\mathbb{E}[\mathbb{I}(I_\ell = I) X_\ell X_\ell']$, $I$ in $\mathcal{I}$, are full rank and in the ASQR case (i) The eigenval-*

*ues of the Gram matrix $\int_{\mathcal{X}} P(x) P'(x)\, dx$ stay bounded away from 0 and infinity when the dimension $K$ of $P(\cdot)$ increases and*

$$\max_{x\in\mathcal{X}} \|P(x)\| = O\left(K^{1/2}\right).$$

*(ii) The sieve $\{P_k, 1 \leq k \leq K\}$ is composed with localized functions, in the sense there is a $c > 0$ such that $P_{k_1}(\cdot) P_{k_2}(\cdot) = 0$ as soon as $|k_2 - k_1| > c/2$ with*

$$\max_{k\leq K} \left\{ \int_{\mathcal{X}} |P_k(x)|\, dx \right\} = O\left(K^{-1/2}\right).$$

*(iii) For some $\eta \in (0,1]$ and $\overline{K}_{1L}$ with $\log \overline{K}_{1L} = O(\log L)$, it holds that*

$$\|P(x) - P(x')\| \leq \overline{K}_{1L} \|x - x'\|^{\eta} \text{ for all } x,x' \text{ of } \mathcal{X}.$$

Assumption R first imposes well conditioned matrices $\mathbb{E}[\mathbb{I}(I_\ell = I) X_\ell X_\ell']$ for the AQR case and $\int_{\mathcal{X}} P(x) P'(x)\, dx$ for the ASQR case. The rest of Assumption R holds for the sieve (4.4) as

$$\max_{x\in\mathcal{X}} \|P(x)\| = O\left(h^{-D_{\mathcal{M}}/2}\right), \quad \max_{k\leq K} \left\{ \int_{\mathcal{X}} |P_k(x)|\, dx \right\} = O\left(h^{-D_{\mathcal{M}}/2}\right)$$

with $K \asymp h^{-1/D_{\mathcal{M}}}$ by Assumption H. Assumption R-(iii) holds when the order $K$ of the sieve (4.4) decreases with a polynomial rate and provided $q(\cdot)$ is Hölder with exponent $\eta$. This allows for cardinal B-splines for which $\eta = 1$, but also for wavelets which are not always differentiable but Hölder with $\eta < 1$, see Daubechies (1992).

## A.2 Uniform consistency rates

The next Theorem deals with uniform consistency of the ASQR procedure.

**Theorem A.1** *Suppose that the private value conditional quantile function $V(\cdot|\cdot)$ is a quantile regression (2.5) or a sieve quantile regression (2.10) with $D_{\mathcal{M}}$ interactions. Then under*

*Assumptions A, H, S and R with* $s \geq D_{\mathcal{M}}/2$ *and*

$$\frac{\log L}{Lh^{2D_{\mathcal{M}}+1+(D_{\mathcal{M}}\vee 1)}} = O\left(1\right),$$

*it holds*

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}} \left|\widehat{V}\left(\alpha|x,I\right) - V\left(\alpha|x,I\right)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2} + h^{s+1}\right),$$
$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}} \left|\widehat{B}\left(\alpha|x,I\right) - B\left(\alpha|x,I\right)\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right) + o\left(h^{s+1}\right).$$

The bandwidth condition used in Theorem A.1 is similar to the one of Theorem 3 and allows an optimal bandwidth of order $(\log L/L)^{1/(2D_{\mathcal{M}}+s+3)}$ provided the smoothness $s$ satisfies

$$s \geq \max\left(\frac{D_{\mathcal{M}}}{2}, D_{\mathcal{M}} - 1\right).$$

Under this condition the uniform consistency rate of the private value conditional quantile estimator is

$$\left(\frac{\log L}{L}\right)^{\frac{s+1}{2s+D_{\mathcal{M}}+3}}$$

which coincides with the GPV optimal minimax uniform consistency rate for the estimation of the private value conditional cdf in the presence of $D_{\mathcal{M}}$ covariates.[1] Theorem A.1 also includes a uniform consistency rate for the bid conditional quantile function estimator which can be used to estimate the bidders' signals and private values.

# References


[1] Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. Chap. 76 in *Handbook of Econometrics,* Vol. **6B**. Elsevier.


[1]GPV consider the pdf but the rate for cdf or quantile can be derived similarly.

[2] Daubechies, I. (1992). *Ten lectures on wavelets.* SIAM.

[3] Härdle, W., G. Kerkyacharian, D. Picard & A. Tsybakov (1998). *Wavelets, approximation and statistical applications.* Springer.

[4] Schumaker, L.L. (2007). *Spline functions: basic theory.* Cambridge University Press.

# Online Appendix B: Notations and intermediary results

We start with additional notations used all along the proof section and some preliminary lemmas which are established in Appendix F. In what follows

$$P(x) = \begin{cases} [1, x']' & \text{in the AQR case } (K = D+1) \\ [P_1(x), \ldots, P_K(x)]' & \text{in the ASQR case} \end{cases}$$

allowing an unified treatment of the two estimators, although the proof focus is on the more difficult ASQR case. Recall that $\|P(x)\| = \left(P(x)' P(x)\right)^{1/2}$ is the standard Euclidean norm and that, under Assumptions R-(i) and H-(ii),

$$\max_{x \in \mathcal{X}} \|P(x)\| = O\left(K^{1/2}\right) = O\left(h^{-D_{\mathcal{M}}/2}\right), \quad \max_{(x,t) \in \mathcal{X} \times [-1,1]} \|P(x,t)\| = O\left(h^{-D_{\mathcal{M}}/2}\right),$$

with $D_{\mathcal{M}} = 0$ in the AQR case. Recall that

$$P(x, ht) = \pi(ht) \otimes P(x), \quad \pi(ht)' = \left[1, ht, \ldots, \frac{(ht)^{s+1}}{(s+1)!}\right]$$

so that the "design" matrix $\mathbb{E}\left[P(x_\ell, ht) P(x_\ell, ht)'\right]$ degenerates asymptotically. To avoid this, consider the change of parameters $\mathsf{b} = Hb$ with $H = \text{Diag}(1, \ldots, h^{s+1}) \otimes \text{Id}_K$,

$$\mathsf{b} = \left[\underbrace{\beta_{0,1}, \ldots, \beta_{0,K}}_{\mathsf{b}_0' = \beta_0'}, \underbrace{h\beta_{1,1}, \ldots, h\beta_{1,K}}_{\mathsf{b}_1' = h\beta_1'}, \ldots, \underbrace{h^{s+1}\beta_{s+1,1}, \ldots, h^{s+1}\beta_{s+1,K}}_{\mathsf{b}_{s+1}' = h^{s+1}\beta_{s+1}}\right] \tag{B.1}$$

so that $P(x_\ell, ht)'\beta = P(x_\ell, t)'\mathsf{b}$. Define accordingly

$$\begin{aligned}
\widehat{\mathsf{R}}(\mathsf{b};\alpha,I) &= \frac{1}{LIh}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_0^1 \rho_a\left(B_{i\ell} - P\left(x_\ell, \frac{a-\alpha}{h}\right)'\mathsf{b}\right)K\left(\frac{a-\alpha}{h}\right)da \\
&= \frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{\mathsf{h}}}^{\frac{1-\alpha}{\mathsf{h}}}\rho_{a+ht}\left(B_{i\ell} - P(x_\ell, t)'\mathsf{b}\right)K(t)\,dt, \\
\overline{\mathsf{R}}(\mathsf{b};\alpha,I) &= \mathbb{E}\left[\widehat{\mathsf{R}}(\mathsf{b};\alpha,I)\right].
\end{aligned}$$

Note that $\mathsf{b} \to \int_{-\frac{\alpha}{\mathsf{h}}}^{\frac{1-\alpha}{\mathsf{h}}}\rho_{a+ht}\left(B_{i\ell} - P(x_\ell, t)'\mathsf{b}\right)K(t)\,dt$ is convex as an integral of convex functions. It follows that $\widehat{\mathsf{R}}(\mathsf{b};\alpha,I)$ and $\overline{\mathsf{R}}(\mathsf{b};\alpha,I)$ have minimizers,

$$\begin{aligned}
\widehat{\mathsf{b}}(\alpha|I) &= \arg\min_{\mathsf{b}}\widehat{\mathsf{R}}(\mathsf{b};\alpha,I) = H\widehat{\beta}(\alpha|I), \\
\overline{\mathsf{b}}(\alpha|I) &= \arg\min_{\mathsf{b}}\overline{\mathsf{R}}(\mathsf{b};\alpha,I),
\end{aligned}$$

which uniqueness will be established in the next section. Set $\overline{b}(\alpha|I) = H^{-1}\overline{\mathsf{b}}(\alpha|I)$ recalling $\overline{b}(\alpha|I) = \left[\overline{\beta}_0(\alpha|I)', \ldots, \overline{\beta}_{s+1}'(\alpha|I)\right]'$ and define $\overline{B}(\alpha|x,I) = P(x)'\overline{\beta}_0(\alpha|I)$,

$$\overline{\gamma}_0(\alpha|I) = \overline{\beta}_0(\alpha|I) + \frac{\alpha\overline{\beta}_1(\alpha|I)}{I-1}, \quad \overline{V}(\alpha|x,I) = P(x)'\overline{\gamma}_0(\alpha|I).$$

By Proposition C.1 and its proof, there exists some $\beta^*(\cdot|\cdot)$ grouping the entries in (2.11) such that

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}|P(x)\beta^*(\alpha|I) - B(\alpha|x,I)| = o\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right) = o\left(h^{s+1}\right).$$

Let $b^*(\cdot|\cdot)$ and $\mathsf{b}^*(\cdot|\cdot) = Hb^*(\cdot|\cdot)$ with

$$\beta^*(\alpha|I)' = \left[\beta_0^*(\alpha|I)', \beta_1^*(\alpha|I)', \ldots, \beta_{s+1}^*(\alpha|I)'\right],$$

$\beta_p^*(\alpha|I) = \left[\beta_k^{(p)}(\alpha|I), 1 \le k \le K\right]$ as in (2.11), $p = 0, \ldots, s+1$.

The next notations deal with the differentiability of the objective functions $\widehat{\mathsf{R}}(\cdot;\alpha,I)$.

Since
$$\frac{\partial\rho_{\alpha+ht}\left(B-P\left(x_{\ell},t\right)'\mathsf{b}\right)}{\partial\mathsf{b}''}=\left\{\mathbb{I}\left(B_{i\ell}\leq P\left(x_{\ell},t\right)'\mathsf{b}\right)-\left(\alpha+ht\right)\right\}P\left(x_{\ell},t\right),$$
almost everywhere, it follows that $\widehat{\mathsf{R}}\left(\cdot;\alpha,I\right)$ is differentiable with

$$\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right)=\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left(I_{\ell}=I\right)\sum_{i=1}^{I_{\ell}}\int_{-\frac{\alpha}{\mathsf{h}}}^{\frac{1-\alpha}{\mathsf{h}}}\left\{\mathbb{I}\left(B_{i\ell}\leq P\left(x_{\ell},t\right)'\mathsf{b}\right)-\left(\alpha+ht\right)\right\}P\left(x_{\ell},t\right)K\left(t\right)dt$$

and $\overline{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right)=\mathbb{E}\left[\widehat{\mathsf{R}}^{(1)}\left(\mathsf{b};\alpha,I\right)\right]$ by the Dominated Convergence Theorem. When $\mathsf{b}=\mathsf{b}^{*}\left(\alpha|I\right)$, $P\left(x,t\right)'\mathsf{b}^{*}\left(\alpha|I\right)=P\left(x,ht\right)'\beta^{*}\left(\alpha|I\right)$ is close to $B\left(\alpha+ht|x,I\right)$, which inverse as a function of $t$ in

$$\mathcal{I}_{\alpha,h}=\left[\underline{I}_{\alpha,h},\overline{I}_{\alpha,h}\right]=\left[-\min\left(1,\frac{\alpha}{h}\right),\min\left(1,\frac{1-\alpha}{h}\right)\right]=\left[-1,1\right]\cap\left[-\frac{\alpha}{h},\frac{1-\alpha}{h}\right]$$

is
$$\frac{G\left(u|x,I\right)-\alpha}{h},\quad u\in\left[B\left(\alpha+h\underline{I}_{\alpha,h}|x,I\right),B\left(\alpha+h\overline{I}_{\alpha,h}|x,I\right)\right].$$

When $h$ is small enough, it will be shown in the proof of Lemma B.1 below that

$$\begin{aligned}\frac{\partial}{\partial t}\left[P\left(x,ht\right)'b^{*}\left(\alpha|I\right)\right]&=h\left[\pi^{(1)}\left(ht\right)\otimes P\left(x\right)\right]'b^{*}\left(\alpha|I\right)\\&=hP\left(x\right)'\beta_{1}^{*}\left(\alpha|I\right)+O\left(h^{2}\right)\end{aligned}$$

uniformly since $\pi^{(1)}\left(ht\right)'=\left[0,1,ht,\ldots,\left(ht\right)^{s}/s!\right]$ and that $P\left(x\right)'\beta_{1}^{*}\left(\alpha|I\right)$ converges uniformly to $B^{(1)}\left(\alpha|x,I\right)$ when $K$ diverges and is therefore positive, so that $P\left(x,t\right)'\mathsf{b}^{*}\left(\alpha|I\right)$ is an increasing function of $t$ in $\mathcal{I}_{\alpha,h}$ for $h$ small enough. Since $\max_{(x,t)\in\mathcal{X}\times[-1,1]}\left\|P\left(x,t\right)\right\|=O\left(h^{-D_{\mathcal{M}}/2}\right)$, $t\rightarrow P\left(x,t\right)'\mathsf{b}$ is also strictly increasing provided $\mathsf{b}$ is close enough to $\mathsf{b}^{*}\left(\alpha|I\right)$.

In such case, it is convenient to redefine $P(x,t)'\mathsf{b}$ as follows[1]

$$\Psi(t|x,\mathsf{b}) = \begin{cases} P\left(x,\overline{I}_{\alpha,h}\right)'\mathsf{b} & t > \overline{I}_{\alpha,h} \\ P(x,t)'\mathsf{b} & t \in \mathcal{I}_{\alpha,h} \\ P\left(x,\underline{I}_{\alpha,h}\right)'\mathsf{b} & t < \underline{I}_{\alpha,h} \end{cases}.$$

When $\Psi(\cdot|x,\mathsf{b})$ has an inverse, define

$$\Phi(u|x,\mathsf{b}) = \begin{cases} \alpha + h\overline{I}_{\alpha,h} & u > \Psi\left(\overline{I}_{\alpha,h}|x,\mathsf{b}\right) \\ \alpha + h\Psi^{-1}(u|x,\mathsf{b}) & u \in \Psi(\mathcal{I}_{\alpha,h}|x,\mathsf{b}) \\ \alpha + h\underline{I}_{\alpha,h} & u < \Psi\left(\underline{I}_{\alpha,h}|x,\mathsf{b}\right) \end{cases},$$

$$\Delta(u|x,\mathsf{b}) = \frac{\Phi(u|x,\mathsf{b}) - \alpha}{h} = \begin{cases} \overline{I}_{\alpha,h} & u > \Psi\left(\overline{I}_{\alpha,h}|x,\mathsf{b}\right) \\ \Psi^{-1}(u|x,\mathsf{b}) & u \in \Psi(\mathcal{I}_{\alpha,h}|x,\mathsf{b}) \\ \underline{I}_{\alpha,h} & u < \Psi\left(\underline{I}_{\alpha,h}|x,\mathsf{b}\right) \end{cases},$$

which is such that, as seen above, the central part of $\Phi(u|x,\mathsf{b}^*(\alpha|I))$ is close to $G(u|x,I)$ when $u$ is in $\Psi(\mathcal{I}_{\alpha,h}|x,\mathsf{b})$. Observe now that, provided $\Psi(\cdot|x,\mathsf{b})$ is increasing and since the support of $K(\cdot)$ is $[-1,1]$

$$\begin{aligned}
&\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \{\mathbb{I}(B_{i\ell} \leq \Psi(t|x_\ell,\mathsf{b})) - (\alpha + ht)\} P(x_\ell,t) K(t)\,dt \\
&= \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \left\{\mathbb{I}\left(\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b}) - \alpha}{h} \leq t\right) - (\alpha + ht)\right\} P(x_\ell,t) K(t)\,dt \\
&= \int_{\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b})-\alpha}{h}}^{\overline{I}_{\alpha,h}} P(x_\ell,t) K(t)\,dt - \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} (\alpha + ht) P(x_\ell,t) K(t)\,dt
\end{aligned}$$

which is differentiable with respect to $\mathsf{b}$, with for $B_{i\ell}$ in $\Psi(\mathcal{I}_{\alpha,h}|x,\mathsf{b})$

$$\frac{\partial \Phi(B_{i\ell}|x_\ell,\mathsf{b})}{\partial \mathsf{b}'} = -\frac{P(x,\Delta(B_{i\ell}|x_\ell,\mathsf{b}))}{\Psi^{(1)}(\Delta(B_{i\ell}|x_\ell,\mathsf{b})|x_\ell,\mathsf{b})/h}\mathbb{I}[B_{i\ell} \in \Psi(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b})].$$

[1] In principle $\Psi(\cdot|\cdot)$ should be denoted $\Psi_{\alpha,h}(\cdot|\cdot)$ to acknowledge that its definition depends upon $\alpha$ and $h$. Instead, $t$ is restricted to lie in $\mathcal{I}_{\alpha,h}$ in the sequel. The same comment applies for the functions $\Psi(\cdot|\cdot)$ and $\Delta(\cdot|\cdot)$ introduced below.

Hence, for $h$ small enough and for $\mathsf{b}$ in the vicinity of $\mathsf{b}^*(\alpha|I)$, $\widehat{\mathsf{R}}(\mathsf{b};\alpha,I)$ and $\overline{\mathsf{R}}(\mathsf{b};\alpha,I)$ are twice continuously differentiable with,

$$
\begin{aligned}
\widehat{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) &= \frac{1}{LIh}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right), I_\ell = I\right] \\
&\qquad \frac{P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)'}{\Psi^{(1)}\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)/h}K\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right), \\
\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) &= \mathbb{E}\left[\widehat{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)\right].
\end{aligned}
$$

The next lemma details some properties of the functions $\Psi(\cdot|x,\mathsf{b})$ and $\Phi(\cdot|x,\mathsf{b})$ that were briefly sketched above. Define

$$
\begin{aligned}
\mathcal{BI}_{\alpha,h} &= \left\{\mathsf{b};\min_{(t,x)\in\mathcal{I}_{\alpha,h}\times\mathcal{X}}\frac{\partial\Psi(t|x,\mathsf{b})}{\partial t}>0\right\}, \\
\underline{\mathcal{BI}}_{\alpha,h} &= \left\{\mathsf{b};\min_{(t,x)\in\mathcal{I}_{\alpha,h}\times\mathcal{X}}\frac{\partial\Psi(t|x,\mathsf{b})}{\partial t}>h/\underline{f},\max_{p=1,\dots,s+1}\left(\frac{\max_{x\in\mathcal{X}}\left|P(x)'\mathsf{b}_p\right|}{h}\right)<\overline{f}\right\},
\end{aligned}
$$

recalling that $\mathsf{b}=\left[\mathsf{b}_0',\dots,\mathsf{b}_{s+1}'\right]'$ and where $\underline{f}$ and $\overline{f}$ will be taken large enough. While $\mathcal{BI}_{\alpha,h}$ is used to bound the first derivative of $\Psi(\cdot|x,\mathsf{b})$ away from 0, $\underline{\mathcal{BI}}_{\alpha,h}$ is used to bound the successive derivatives $\Psi^{(p)}(\cdot|x,\mathsf{b})$, $p=1,\dots,s+1$, away from infinity. As made possible by Lemma B.1-(i), below, an Euclidean ball $\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)$ with a small enough constant $C>0$ will be considered instead of the sets $\mathcal{BI}_{\alpha,h}$ and $\underline{\mathcal{BI}}_{\alpha,h}$.

**Lemma B.1** *Suppose Assumptions A and S hold with* $\max_{x\in\mathcal{X}}\|P(x)\|=O\left(K^{1/2}\right)$*,* $K=h^{1/D_{\mathcal{M}}}$ *that* $\underline{f}$ *and* $\overline{f}$ *are large enough. Then, $h$ small enough and all $I$ in $\mathcal{I}$,*

*i.* $\mathsf{b}^*(\alpha|I)$ *belongs to* $\underline{\mathcal{BI}}_{\alpha,h}\subset\mathcal{BI}_{\alpha,h}$ *and for $C$ small enough* $\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)$ *is a subset of* $\underline{\mathcal{BI}}_{\alpha,h}$*, for all $\alpha$ in* $[0,1]$.

*ii. For all* $\mathsf{b}$ *in* $\mathcal{BI}_{\alpha,h}$ *and all* $u$ *in* $\Psi\left(\mathcal{I}_{\alpha,h}|x,\mathsf{b}\right)$

$$\begin{aligned}
\frac{\partial\Phi\left(u|x,\mathsf{b}\right)}{\partial\mathsf{b}'} &= -\frac{P\left(x,\Delta\left(u|x,\mathsf{b}\right)\right)}{\Psi\left(\Delta\left(u|x,\mathsf{b}\right)|x,\mathsf{b}\right)/h},\\
\frac{\partial\Phi\left(u|x,\mathsf{b}\right)}{\partial u} &= \frac{1}{\Psi\left(\Delta\left(u|x,\mathsf{b}\right)|x,\mathsf{b}\right)/h}.
\end{aligned}$$

*iii. It holds that*

$$\begin{aligned}
&\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)-B\left(\alpha+ht|x,I\right)\right| = o\left(h^{s+1}\right),\\
&\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\Bigg|\alpha\left(B\left(\alpha+ht|x,I\right)-\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)\right)\\
&\qquad\qquad\qquad\qquad -\frac{(ht)^{s+2}}{(s+2)!}\alpha B^{(s+2)}\left(\alpha|I\right)\Bigg| = o\left(h^{s+2}\right),
\end{aligned}$$

*and, recalling* $\mathsf{b}_1^*\left(\alpha|I\right)=h\beta_1^*\left(\alpha|I\right)$

$$\begin{aligned}
\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\alpha\beta_1^*\left(\alpha|I\right)-\alpha B^{(1)}\left(\alpha|x,I\right)\right| &= o\left(h^{s+1}\right),\\
\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|\Phi\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)-G\left(u|x,I\right)\right| &= o\left(h^{s+1}\right).
\end{aligned}$$

*iv. There is a* $C>0$ *such that for any* $\mathsf{b}_0$ *and* $\mathsf{b}_1$ *in* $\underline{\mathcal{BI}}_{\alpha,h}$ *and all* $\alpha$ *in* $[0,1]$

$$\begin{aligned}
&\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi\left(t|x,\mathsf{b}_1\right)-\Psi\left(t|x,\mathsf{b}_0\right)\right|,\\
&\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_0\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]}\left|\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)\right|,\\
&\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_0\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]}\left|\frac{\partial\Phi}{\partial u}\left(u|x,\mathsf{b}_1\right)-\frac{\partial\Phi}{\partial u}\left(u|x,\mathsf{b}_0\right)\right|,\\
&\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_0\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]}\left|\Psi^{(1)}\left(\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right)-\Psi^{(1)}\left(\Delta\left(u|x_\ell,\mathsf{b}_0\right)|x_\ell,\mathsf{b}_0\right)\right|,
\end{aligned}$$

*are all smaller or equal to* $Ch^{-D_{\mathcal{M}}/2}\left\|\mathsf{b}_1-\mathsf{b}_0\right\|$.

Let $\Omega_h(\alpha)$, $\Omega(0)$, $\Omega(1)$, $\Omega = \Omega(0) + \Omega(1)$ and $\Omega_{1h}(\alpha)$ be the $(s+2) \times (s+2)$ matrices

$$\begin{aligned}
\Omega_h(\alpha) &= \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \pi(t)\pi(t)' K(t)\,dt = \left[\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} t^{p_1+p_2} K(t)\,dt, 0 \leq p_1, p_2 \leq s+1\right], \\
\Omega(0) &= \int_{-1}^{0} \pi(t)\pi(t)' K(t)\,dt, \quad \Omega(1) = \int_{0}^{1} \pi(t)\pi(t)' K(t)\,dt, \\
\Omega_{1h}(\alpha) &= \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} t\pi(t)\pi(t)' K(t)\,dt,
\end{aligned}$$

While $\Omega_h(\alpha) \preceq \Omega$ for all $\alpha$ and $h$, it holds that for $h$ small enough $\Omega_h(\alpha) \succeq \Omega(0)$ for all $\alpha$ in $[0, 1/2]$ and $\Omega_h(\alpha) \succeq \Omega(1)$ for all $\alpha$ in $[1/2, 1]$.

**Lemma B.2** *Suppose Assumptions A, R-(i) and S hold, that $\underline{f}$ and $\overline{f}$ are large enough. Then, for $K^{-1/D_{\mathcal{M}}} = O(h)$, $h$ small enough, all $I$ in $\mathcal{I}$, and any $C > 0$ small enough, (i) It holds that $\overline{\mathsf{R}}^{(2)}(\cdot; \alpha, I)$ is continuously differentiable over $\mathcal{B}\left(\mathsf{b}^*(\alpha|I), Ch^{D_{\mathcal{M}}/2+1}\right)$ with*

$$\max_{\alpha \in [0,1]} \max_{\mathsf{b}_1, \mathsf{b}_0 \in \mathcal{B}\left(\mathsf{b}^*(\alpha|I), Ch^{D_{\mathcal{M}}/2+1}\right)} \frac{\left\|\overline{\mathsf{R}}^{(2)}(\mathsf{b}_1; \alpha, I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b}_0; \alpha, I)\right\|}{\|\mathsf{b}_1 - \mathsf{b}_0\| / (\alpha(1-\alpha) + h)} = O\left(h^{-D_{\mathcal{M}}/2}\right).$$

*(ii) The eigenvalues of $\overline{\mathsf{R}}^{(2)}[\mathsf{b}^*(\alpha|I); \alpha, I]$ belongs to $[1/C, C]$ for a large enough $C$, for all $\alpha$ in $[0, 1]$ and $h$ small enough with*

$$\begin{aligned}
\max_{\alpha \in [0,1]} \Bigg\| \overline{\mathsf{R}}^{(2)}[\mathsf{b}^*(\alpha|I); \alpha, I] &- \Omega_h(\alpha) \otimes \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = 1) P(x_\ell) P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] \\
&+ h\Omega_{1h}(\alpha) \otimes \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = 1) B^{(2)}(\alpha|x_\ell, I_\ell) P(x_\ell) P(x_\ell)'}{\left(B^{(1)}(\alpha|x_\ell, I_\ell)\right)^2}\right] \Bigg\| = o(h).
\end{aligned}$$

Lemma B.2-(i) yields, for any $C > 0$,

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}(\mathsf{b}^*(\alpha|I),Ch^{s+1})} \left\|\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b}^*(\alpha|I);\alpha,I)\right\| = O\left(h^{s-D_{\mathcal{M}}/2}\right)$$
$$\text{if } h^s = o\left(h^{D_{\mathcal{M}}/2}\right),$$
$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),C\left(\frac{\log L}{L}(\alpha(1-\alpha)+h)\right)^{1/2}\right)} \frac{\left\|\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b}^*(\alpha|I);\alpha,I)\right\|}{\left(\frac{\log L}{L(\alpha(1-\alpha)+h)}\right)^{1/2}} = O\left(h^{-D_{\mathcal{M}}/2}\right)$$
$$\text{if } \left(\frac{\log L}{L}\right)^{1/2} = o\left(h^{D_{\mathcal{M}}/2+1}\right).$$

It then follows that the eigenvalues of $\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)$ stays bounded away from 0 and infinity uniformly in $\alpha$ and in $\mathsf{b}$ in the two neighborhoods considered above, under the corresponding bandwidth assumption.

The two next Lemmas study the first and second derivatives of $\widehat{\mathsf{R}}(\cdot;\alpha,I)$ in a shrinking vicinity of $\mathsf{b}^*(\alpha|I)$. In particular, Lemma B.3 implies that $\widehat{\mathsf{R}}(\cdot;\alpha,I)$ is strictly convex over such a vicinity with a probability tending to 1.

**Lemma B.3** *Suppose Assumptions A, R-(i,ii) and S hold, and* $\log L/\left(Lh^{D_{\mathcal{M}}+1}\right) = o(1)$. *Then, for any* $C > 0$ *small enough,*

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)} \left\|\widehat{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)\right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right)$$

**Lemma B.4** *Suppose Assumptions A, R-(i,ii) and S hold, and* $\log L/\left(Lh^{D_{\mathcal{M}}+1}\right) = o(1)$. *Then, for any* $C > 0$,

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)} \left\|\frac{\widehat{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)}{(h+\alpha(1-\alpha))^{1/2}}\right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right).$$

Since $\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) = 0$ and assuming $h^{s+1} = O\left(h^{D_{\mathcal{M}}/2+1}\right)$, $\sup_{\alpha\in[0,1]}\left\|\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right\| =$

$o\left(h^{s+1}\right)$ as established in (C.3), it holds that

$$\max_{\alpha\in[0,1]}\left\|\frac{\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)}{\left(h+\alpha\left(1-\alpha\right)\right)^{1/2}}\right\|=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right).$$

The next Lemma studies the leading term $\widehat{\mathsf{e}}\left(\alpha|I\right)$ of $\widehat{\mathsf{b}}\left(\alpha|I\right)-\overline{\mathsf{b}}\left(\alpha|I\right)$,

$$\widehat{\mathsf{e}}\left(\alpha|I\right)=-\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\right]^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)$$

see Theorem D.1 below. Note that $\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)$ is not necessarily defined and invertible unless $h^{s+1}=O\left(h^{D_{\mathcal{M}}/2+1}\right)$ and $\sup_{\alpha\in[0,1]}\left\|\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^{*}\left(\alpha|I\right)\right\|=o\left(h^{s+1}\right)$ as therefore assumed and established in the proof of Theorem C.4 below, see (C.3).

**Lemma B.5** *Suppose Assumptions A, H, R and S hold, and* $1/\left(Lh^{D_{\mathcal{M}}+1}\right)=o(1)$, $s\geq D_{\mathcal{M}}/2$ *and* $\sup_{\alpha\in[0,1]}\left\|\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^{*}\left(\alpha|I\right)\right\|=o\left(h^{s+1}\right)$*. Then (i) uniformly in* $(\alpha,x)$ *in* $[0,1]\times\mathcal{X}$

$$\operatorname{Var}\left[P\left(x\right)'\widehat{\mathsf{e}}_0\left(\alpha|I\right)\right]=O\left(\frac{1}{Lh^{D_{\mathcal{M}}}}\right)$$

*and* $\operatorname{Var}\left[P\left(x\right)'\widehat{\mathsf{e}}_1\left(\alpha|I\right)/h\right]=O\left(\frac{1}{Lh^{D_{\mathcal{M}}+1}}\right)$ *with* $\operatorname{Var}\left[\widehat{\mathsf{e}}_1\left(\alpha|I\right)/h\right]$ *having the expansion*

$$v_h^2\left(\alpha\right)\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)'P\left(x_\ell\right)}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]\mathbb{E}\left[\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)'P\left(x_\ell\right)\right]\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)'P\left(x_\ell\right)}{B^{(1)}\left(\alpha|x_\ell,I_\ell\right)}\right]+o(1).$$

*(ii) It also holds*

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\widehat{\mathsf{e}}_0\left(\alpha|I\right)\right|=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right),$$
$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\frac{\widehat{\mathsf{e}}_1\left(\alpha|I\right)}{h}\right|=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right).$$

# Online Appendix C: Asymptotic bias

Our bias results for the bid quantile function are based on the next Proposition, which states bid implications of Assumption S.

**Proposition C.1** *Assume the approximation property S holds. Suppose that $V(\alpha|x,I)$ is a $(s+1)$th continuously differentiable function over $[0,1]\times\mathcal{X}$ satisfying,*

$$\inf_{(\alpha,x)\in[0,1]\times\mathcal{X}} V^{(1)}(\alpha|x,I)>0 \text{ and } \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} V^{(1)}(\alpha|x,I)<\infty.$$

*Then, for $B(\alpha|x,I)$ as in (2.3) and sieve coefficients $\{\gamma_k(\alpha|I), 1\le k\le K\}$ of $V(\alpha|x,I)$ as in Property S*

*i.* $\min_{(\alpha,x)\in[0,1]\times\mathcal{X}} B^{(1)}(\alpha|x,I)>0$, $\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} B^{(1)}(\alpha|x,I)<\infty$ *and $B(\alpha|x,I)$ is $(s+2)$th continuously differentiable over $(0,1]$ with*

$$\lim_{\alpha\to 0}\sup_{(x,I)\in\mathcal{X}\times\mathcal{I}}\left|\alpha B^{(s+2)}(\alpha|x,I)\right|=0.$$

*ii. The coefficients $\{\beta_k(\alpha|I), 1\le k\le K\}$ from (2.11) are $(s+1)$th continuously differentiable and satisfy*

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B(\alpha|x,I)-\sum_{k=1}^{K}\beta_k(\alpha|I)P_k(x)\right|=o\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right),$$
$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B^{(p)}(\alpha|x,I)-\sum_{k=1}^{K}\beta_k^{(p)}(\alpha)P_k(x)\right|=o(1),\quad p=1,\ldots,s+1.$$

*iii. Moreover $\alpha\beta_k^{(1)}(\alpha)=(I-1)[\gamma_k(\alpha|I)-\beta_k(\alpha)]$ and is therefore $(s+1)$th continuously* *differentiable for all* $1 \leq k \leq K$. *In addition*

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| \alpha B^{(1)}(\alpha|x,I) - \sum_{k=1}^{K} \alpha\beta_k^{(1)}(\alpha|x,I) P_k(x) \right| = o\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right),$$
$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left| \frac{\partial^p \left[\alpha B^{(1)}(\alpha|x,I)\right]}{\partial\alpha^p} - \sum_{k=1}^{K} \frac{\partial^p \left[\alpha\beta_k^{(1)}(\alpha|x,I)\right]}{\partial\alpha^p} P_k(x) \right| = o(1), \quad p = 1, \ldots, s+1.$$

**Proof of Proposition C.1.** By (2.3), $B(\alpha|x,I) = (I-1)\int_0^1 u^{I-2} V(\alpha u|x,I)\,du$, so that $B^{(1)}(\alpha|x,I) = (I-1)\int_0^1 u^{I-1} V^{(1)}(\alpha u|x,I)\,du$ which implies the two first statements in (i) about lower and upper bounds for $B^{(1)}(\alpha|x,I)$ and that $B(\cdot|\cdot,I)$ is $(s+1)$th continuously differentiable. That $B(\cdot|x,I)$ is $(s+2)$th continuously differentiable over $(0,1]$ follows from its integral expression (2.3). Observe now that for $p = 1, \ldots, s+2$

$$\frac{\partial^p \left[\alpha B(\alpha|x,I)\right]}{\partial\alpha^p} = \alpha B^{(p)}(\alpha|x,I) + pB^{(p-1)}(\alpha|x,I)$$

with, for $p = 1, \ldots, s+1$

$$\begin{aligned}
B^{(p)}(\alpha|x,I) &= (I-1)\int_0^1 u^{I-2+p} V^{(p)}(\alpha u|x,I)\,du = \frac{I-1}{\alpha^{I-1+p}}\int_0^\alpha t^{I-2+p} V^{(p)}(t|x,I)\,dt \\
B^{(p+1)}(\alpha|x,I) &= -\frac{(I-1)(I-1+p)}{\alpha^{I+p}}\int_0^\alpha t^{I-2+p} V^{(p)}(t|x,I)\,dt + \frac{(I-1)V^{(p)}(\alpha|x,I)}{\alpha} \\
&= -\frac{I-1+p}{\alpha}B^{(p)}(\alpha|x,I) + \frac{(I-1)V^{(p)}(\alpha|x,I)}{\alpha}.
\end{aligned}$$

Hence, when $\alpha$ goes to 0

$$\begin{aligned}
\alpha B^{(s+2)}(\alpha|x,I) &= -(I+s)B^{(s+1)}(0|x,I) + (I-1)V^{(s+1)}(0|x,I) + o(1) \\
&= -(I+s)(I-1)\int_0^1 u^{I+s-1} V^{(s+1)}(0|x,I)\,du + (I-1)V^{(s+1)}(0|x,I) + o(1) \\
&= o(1)
\end{aligned}$$

uniformly on $x$.

For (ii), consider a sequence of $\{\gamma_k(\alpha|I), k \leq K\}$ approximating $V(\alpha|x,I)$ and its derivatives as in Property S. For $\{\beta_k(\alpha|I), k \leq K\}$ as in (2.11)

$$\beta_k^{(p)}(\alpha|I) = (I-1)\int_0^1 u^{I+p-2}\gamma_k^{(p)}(\alpha u|I)\,du, \quad p = 0,\ldots,s+1$$

and

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B^{(p)}(\alpha|x,I) - \sum_{k=1}^{K}\beta_k^{(p)}(\alpha|I)P_k(x)\right| \\
&= \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|(I-1)\int_0^1 u^{I+p-2}\left(V^{(p)}(\alpha u|x,I) - \sum_{k=1}^{K}\gamma_k^{(p)}(\alpha u|I)P_k(x)\right)du\right| \\
&\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|V^{(p)}(\alpha|x,I) - \sum_{k=1}^{K}\gamma_k^{(p)}(\alpha|I)P_k(x)\right|
\end{aligned}$$

which gives the sieve approximation result for $B(\alpha|x,I)$ in (ii). Now, for $\alpha B^{(1)}(\alpha|x,I)$, observe that $\alpha B^{(1)}(\alpha|x,I) = (I-1)[V(\alpha|x,I) - B(\alpha|x,I)]$ and

$$\begin{aligned}
\alpha\beta_k^{(1)}(\alpha|I) &= \alpha \times \left(-\frac{(I-1)^2}{\alpha^I}\int_0^1 t^{I-2}\gamma_k(t|I)\,dt + \frac{I-1}{\alpha}\gamma_k(\alpha|I)\right) \\
&= (I-1)[\gamma_k(\alpha|I) - \beta_k(\alpha|I)].
\end{aligned}$$

It follows

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{\partial^p\left[\alpha B^{(1)}(\alpha|x,I)\right]}{\partial\alpha^p} - \sum_{k=1}^{K}\frac{\partial^p\left[\alpha\beta_k^{(1)}(\alpha|I)\right]}{\partial\alpha^p}P_k(x)\right| \\
&\leq (I-1)\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|V^{(p)}(\alpha|x,I) - \sum_{k=1}^{K}\gamma_k^{(p)}(\alpha|I)P_k(x)\right| \\
&+ (I-1)\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|B^{(p)}(\alpha|x,I) - \sum_{k=1}^{K}\beta_k^{(p)}(\alpha|I)P_k(x)\right| \\
&\leq 2(I-1)\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|V^{(p)}(\alpha|x,I) - \sum_{k=1}^{K}\gamma_k^{(p)}(\alpha|I)P_k(x)\right|
\end{aligned}$$

which gives the approximation result for $\alpha B^{(1)}(\alpha|x,I)$ in (iii). $\square$

The study of the bias $\overline{V}(\alpha|x,I) - V(\alpha|x,I)$ and $\overline{B}(\alpha|x,I) - B(\alpha|x,I)$ is based on the following Lemma which is a consequence of the Kantorovitch-Newton Theorem, see e.g. Gragg and Tapia (1974).

**Lemma C.2** *Let $\mathcal{F}(\cdot) : \mathbb{R}^D \to \mathbb{R}$ be a function. Suppose that there is a $\mathbf{x}^* \in \mathbb{R}^D$ and some real numbers $\epsilon > 0$ and $C_0 > 0$ such that $\mathcal{F}(\cdot)$ is twice differentiable on $\mathcal{B}(\mathbf{x}^*, 2C_0\epsilon) = \left\{x \in \mathbb{R}^D; \|x - \mathbf{x}^*\| < 2C_0\epsilon\right\}$. If, in addition,*

*i.* $\left\|\mathcal{F}^{(1)}(\mathbf{x}^*)\right\| \leq \epsilon$ *and* $\left\|\left[\mathcal{F}^{(2)}(\mathbf{x}^*)\right]^{-1}\right\| \leq C_0$*;*

*ii. There is a $C_1 > 0$ such that $\left\|\mathcal{F}^{(2)}(x) - \mathcal{F}^{(2)}(x')\right\| \leq C_1 \|x - x'\|$ for all $x, x' \in \mathcal{B}(\mathbf{x}^*, 2C_0\epsilon)$;*

*iii.* $C_0^2 C_1 \epsilon \leq 1/2$.

*Then there is a unique $\overline{\mathbf{x}}$ such that $\|\overline{\mathbf{x}} - \mathbf{x}^*\| < 2C_0\epsilon$ and $\mathcal{F}^{(1)}(\overline{\mathbf{x}}) = 0$.*

The next lemma, established in Appendix F, will be used at the end of the proof of Theorem C.4 below.

**Lemma C.3** *Suppose Assumptions A, S and R-(ii). Then the $\ell_1$ norm of the columns of the matrix*

$$A_{\alpha,h} = \mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P(x_\ell, t) P(x_\ell, t)' K(t)\, dt}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right]$$

*are bounded independently of L and $\alpha$. That is, if $A_{\alpha,h} = [A_{\alpha,h}(j_1, j_2), 1 \leq j_1, j_2 \leq (s+1)K]$,*

$$\max_L \max_{\alpha \in [0,1]} \max_{1 \leq j_1 \leq (s+1)K} \sum_{j_2=1}^{(s+1)K} |A_{\alpha,h}(j_1, j_2)| < \infty.$$

In the next theorem,

$$\mathsf{bias}_h(\alpha|I) = \mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P(x_\ell,t)\,P(x_\ell,t)'\,K(t)\,dt}{B^{(1)}(\alpha|x_\ell,I_\ell)}\right]$$
$$\times\,\mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)\,B^{(s+2)}(\alpha|x_\ell,I_\ell)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} t^{s+2}P(x_\ell,t)\,K(t)\,dt}{(s+2)!B^{(1)}(\alpha|x_\ell,I_\ell)}\right],$$

and

$$\mathsf{bias}_h(\alpha|I) = \left[\mathsf{bias}_{0h}(\alpha|I)', \ldots, \mathsf{bias}_{s+1,h}(\alpha|I)'\right]$$

where the subvectors $\mathsf{bias}_{ph}(\alpha|I)$ are of dimension $K$. While $\mathsf{bias}_h(\alpha|I)$ may not exist for $\alpha = 0$, the function $\mathsf{Bias}_h(\alpha|I) = \alpha\mathsf{bias}_h(\alpha|I)$ in (4.8) can be set to 0 when $\alpha = 0$ by Proposition C.1-(i).

**Theorem C.4** *Suppose that Assumptions A, H and R hold with $s \geq D_{\mathcal{M}}/2$. Then, for $h$ small enough $\overline{\mathsf{b}}(\alpha|I) = \arg\min_{\mathsf{b}} \overline{\mathsf{R}}(\mathsf{b};\alpha,I)$ is unique for all $\alpha$ in $[0,1]$ and*

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\overline{V}(\alpha|x,I) - V(\alpha|x,I) - \frac{h^{s+1}P(x)'\,\alpha\mathsf{bias}_{1h}(\alpha|I)}{I-1}\right| = o\left(h^{s+1}\right)$$

*with* $\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|P(x)'\,\alpha\mathsf{bias}_{1h}(\alpha|I)\right| = O(1)$. *Moreover*

$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\overline{B}(\alpha|x,I) - B(\alpha|x,I)\right| = o\left(h^{s+1}\right),$$
$$\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left|\overline{B}^{(1)}(\alpha|x,I) - B^{(1)}(\alpha|x,I)\right| = o\left(h^{s}\right).$$

The proof of Theorem C.4 establishes that $\sup_{\alpha\in[0,1]}\left\|\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right\| = o(h^{s+1})$, see (C.3), an intermediary result which will be used all along the proof. If $D_{\mathcal{M}}/2 \leq s$,

$\log L / \left( L h^{D_{\mathcal{M}}+1} \right) = o(1)$ and by Lemma B.3 and a second order Taylor expansion

$$\sup_{\alpha \in [0,1]} \sup_{\mathsf{b} \in \mathcal{B}\left(\overline{\mathsf{b}}(\alpha|I), C h^{s+1}\right)} \left| h^{-2(s+1)} \left\{ \widehat{\mathsf{R}}(\mathsf{b}; \alpha, I) - \widehat{\mathsf{R}}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right) - \left(\mathsf{b} - \overline{\mathsf{b}}(\alpha|I)\right)' \widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right) \right\} \right.$$
$$\left. - \frac{h^{-2(s+1)}}{2} \left(\mathsf{b} - \overline{\mathsf{b}}(\alpha|I)\right)' \overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right) \left(\mathsf{b} - \overline{\mathsf{b}}(\alpha|I)\right) \right| = o_{\mathbb{P}}(1).$$

Then by Lemma B.2 and the Argmax Theorem $\widehat{\mathsf{R}}(\cdot; \alpha, I)$ has a unique minimizer over $\mathsf{b} \in \mathcal{B}\left(\overline{\mathsf{b}}(\alpha|I), C h^{s+1}\right)$ for each $\alpha$, with a probability tending to 1. Since $\widehat{\mathsf{R}}(\cdot; \alpha, I)$ is convex a local minimum is also a global one. This implies that the AQR or ASQR estimators $\widehat{b}(\alpha|I) = H^{-1} \widehat{\mathsf{b}}(\alpha|I)$ are unique for all $\alpha$ in $[0,1]$ with a probability tending to 1.

**Proof of Theorem C.4.** Consider (ii) and (iii), the proof of (i) being similar as detailed below. The proof works by establishing that there is a solution of the first-order condition in a open ball where $\overline{\mathsf{R}}(\mathsf{b}; \alpha, I)$ is strictly convex by checking the conditions of Lemma C.2, which will also gives the rate stated in the Theorem and the uniqueness of $\overline{\mathsf{b}}(\alpha|I)$. It is first claimed that

$$\max_{(\alpha, I) \in [0,1] \times \mathcal{I}} \left\| \overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^*(\alpha|I); \alpha, I\right) \right\| = \epsilon_L \text{ with} \tag{C.1}$$
$$\epsilon_L = O\left( \max_{(\alpha, x) \in [0,1] \times \mathcal{X}} \max_{t \in \mathcal{I}_{\alpha,h}} \left| \Psi\left(t|x, \mathsf{b}^*(\alpha|I)\right) - B(\alpha + ht|x, I) \right| \right) = o\left(h^{s+1}\right),$$

where $\epsilon_L = o\left(h^{s+1}\right)$ follows from Lemma B.1-(iii). To see that (C.1) holds, observe that

$$\left\| \overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^*(\alpha|I); \alpha, I\right) \right\| = \max_{\theta; \theta'\theta = 1} \left| \theta' \overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^*(\alpha|I); \alpha, I\right) \right|. \tag{C.2}$$

But uniformly in $\alpha \in [0,1]$ and by Assumption R-(i), Lemma B.1-(iii),

$$
\begin{aligned}
&\left|\theta' \overline{\mathsf{R}}^{(1)}\left(\mathsf{b}^*\left(\alpha|I\right);\alpha,I\right)\right| \\
&= \mathbb{E}\left[\mathbb{I}\left(I_\ell = I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \left\{G\left(P\left(x_\ell,t\right)\mathsf{b}^*\left(\alpha|I\right)|x_\ell,I_\ell\right) - G\left(B\left(\alpha+ht|x,I\right)|x_\ell,I_\ell\right)\right\}\right. \\
&\qquad\qquad \left.\theta'\left(P\left(x_\ell\right)\otimes\pi\left(t\right)\right)K\left(t\right)dt\right] \\
&\leq C\epsilon_L \mathbb{E}^{1/2}\left[\int_{-1}^{1}\left(\theta'\left(P\left(x_\ell\right)\otimes\pi\left(t\right)\right)\right)^2 dt\right] \leq C\epsilon_L\left(\theta'\theta\right)^{1/2} = C\epsilon_L.
\end{aligned}
$$

Hence (C.1) holds, which is the first part of Condition (i) in Lemma C.2. The second part of Condition (i) follows from Lemma B.2-(ii) which ensures that there is a $C_0 > 0$ such that, for $L$ large enough,

$$
\sup_{(\alpha,I)\in[0,1]\times\mathcal{I}}\left\|\left[\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}^*\left(\alpha|I\right);\alpha,I\right)\right]^{-1}\right\| \leq C_0
$$

Note that $s \geq D_{\mathcal{M}}/2$ and $\epsilon_L = o\left(h^{s+1}\right)$ gives that

$$
\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),2C_0\epsilon_L\right) \subset \mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),Ch^{D_{\mathcal{M}}/2+1}\right)
$$

for all $C_0, C > 0$ provided $L$ is large enough, for all $\alpha$ and all $I$. Condition (ii) in Lemma C.2 follows from Lemma B.2-(i) which ensures that for $C_{1L} = O\left(h^{D_{\mathcal{M}}/2+1}\right)$,

$$
\left\|\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}_1;\alpha,I\right) - \overline{\mathsf{R}}^{(2)}\left(\mathsf{b}_0;\alpha,I\right)\right\| \leq C_{1L}\left\|\mathsf{b}_1 - \mathsf{b}_0\right\|
$$

for all $\mathsf{b}_1$, $\mathsf{b}_0$ in $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),2C_0\epsilon_L\right)$ and all $\alpha$, $I$. For condition (iii) in Lemma C.2, $\epsilon_L = o\left(h^{s+1}\right)$ and $s \geq D_{\mathcal{M}}/2$ implies $C_0^2 C_{1L}\epsilon_L = o\left(h^{s-D_{\mathcal{M}}/2}\right) = o\left(1\right) < 1/2$ for $L$ large enough. Hence Lemma C.2 ensures that, for $L$ large enough, all $\alpha$ and all $I$, there is a unique $\overline{\mathsf{b}}\left(\alpha|I\right)$ in $\mathcal{B}\left(\mathsf{b}^*\left(\alpha|I\right),2C_0\epsilon_L\right)$ such that

$$
\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right) = 0
$$

and is therefore the unique minimizer of $\overline{\mathsf{R}}(\cdot;\alpha,I)$ over $\mathcal{B}(\mathsf{b}^*(\alpha|I),2C_0\epsilon_L)$. Since the convex function $\overline{\mathsf{R}}(\cdot;\alpha,I)$ cannot have several local minimizers, $\overline{\mathsf{b}}(\alpha|I)$ is also the unique global minimizer of $\overline{\mathsf{R}}(\cdot;\alpha,I)$. Since $\epsilon_L = o(h^{s+1})$, it follows that

$$\sup_{(\alpha,I)\in[0,1]\times\mathcal{I}} \left\|\overline{\mathsf{b}}(\alpha|I)-\mathsf{b}^*(\alpha|I)\right\| = o\left(h^{s+1}\right). \tag{C.3}$$

Consider now $\alpha\overline{\mathsf{b}}(\alpha|I)-\alpha\mathsf{b}^*(\alpha|I)$. Define

$$\overline{g}(\alpha|t,x,I) = \int_0^1 g\left(\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)+u\left(B(\alpha+ht|x,I)-\Psi(t|x,\mathsf{b}^*(\alpha|I))\right)|t,x,I\right)du$$

which is such that, uniformly in $\alpha$ in $[3h,1]$, $x$ in $\mathcal{X}$ and $t$ in $[-1,3/4]$

$$\begin{aligned}
\overline{g}(\alpha|t,x,I) &= \int_0^1 g\left(\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)+u\left(B(\alpha+ht|x,I)-\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)\right)|t,x,I\right)du\\
&= \int_0^1 g\left(B(\alpha+ht|x,I)+o\left(h^{s+1-D_{\mathcal{M}}/2}\right)|t,x,I\right)du\\
&\geq (1+o(1))\max_{y\in[B(2h|x,I),B(1-2h|x,I)]} g(y|x,I)\geq C''>0
\end{aligned}$$

by Lemma B.1-(iii,iv), (C.3), $o\left(h^{s+1-D_{\mathcal{M}}/2}\right)=o(h)$ and Proposition C.1-(i). Now $\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) = 0$ gives

$$\begin{aligned}
0 &= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left\{G\left[\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)|x,I\right]-(\alpha+ht)\right\}P(x,t)K(t)\,dt\right)f(x,I)\,dx\\
&= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left\{G\left[\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)|x,I\right]-G\left[B(\alpha+ht|x,I)|x,I\right]\right\}P(x,t)K(t)\,dt\right)f(x,I)\,dx\\
&= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}(\alpha|t,x,I)\left\{\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)-B(\alpha+ht|x,I)\right\}P(x,t)K(t)\,dt\right)f(x,I)\,dx\\
&= \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}(\alpha|t,x,I)\left\{\Psi\left(t|x,\overline{\mathsf{b}}(\alpha|I)\right)-\Psi(t|x,\mathsf{b}^*(\alpha|I))\right\}P(x,t)K(t)\,dt\right)f(x,I)\,dx\\
&+ \int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}(\alpha|t,x,I)\left\{\Psi(t|x,\mathsf{b}^*(\alpha|I))-B(\alpha+ht|x,I)\right\}P(x,t)K(t)\,dt\right)f(x,I)\,dx.
\end{aligned}$$

Since $\left\{\Psi\left(t|x,\overline{\mathsf{b}}\left(\alpha|I\right)\right)-\Psi\left(t|x,\mathsf{b}^{*}\left(\alpha|I\right)\right)\right\}P\left(x,t\right)=P\left(x,t\right)P\left(x,t\right)'\left(\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^{*}\left(\alpha|I\right)\right)$, by Assumption R-(i), and because $\overline{g}\left(\alpha|t,x,I\right)$, $f\left(x,I\right)$ are bounded away from 0 and infinity

$$\alpha\left(\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^{*}\left(\alpha|I\right)\right)=\left[\int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}\left(\alpha|t,x,I\right)P\left(x,t\right)P\left(x,t\right)'K\left(t\right)dt\right)f\left(x,I\right)dx\right]^{-1}$$
$$\times\int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\overline{g}\left(\alpha|t,x,I\right)\left\{\frac{\left(ht\right)^{s+2}}{\left(s+2\right)!}\alpha B^{\left(s+2\right)}\left(\alpha|x,I\right)+o\left(h^{s+2}\right)\right\}P\left(x,t\right)K\left(t\right)dt\right)f\left(x,I\right)dx$$

uniformly in $\alpha$ in $[0,1]$ by Lemma B.1-(iii). By Assumption R-(ii) which implies in particular $\left\|\int\left(\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left|P\left(x,t\right)\right|K\left(t\right)dt\right)dx\right\|=O\left(1\right)$, it follows

$$\begin{aligned}&\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^{*}\left(\alpha|I\right)\\&\quad=o\left(h^{s+1}\right)\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_{\ell}=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}P\left(x_{\ell},t\right)P\left(x_{\ell},t\right)'K\left(t\right)dt}{B^{\left(1\right)}\left(\alpha|x_{\ell},I_{\ell}\right)}\right]\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left|P\left(x_{\ell},t\right)\right|K\left(t\right)dt\right],\end{aligned}$$

$$\begin{aligned}&\alpha\left(\overline{\mathsf{b}}\left(\alpha|I\right)-\mathsf{b}^{*}\left(\alpha|I\right)\right)=h^{s+2}\alpha\mathsf{bias}_{h}\left(\alpha|I\right)\\&\quad+o\left(h^{s+2}\right)\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_{\ell}=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}P\left(x_{\ell},t\right)P\left(x_{\ell},t\right)'K\left(t\right)dt}{B^{\left(1\right)}\left(\alpha|x_{\ell},I_{\ell}\right)}\right]\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left|P\left(x_{\ell},t\right)\right|K\left(t\right)dt\right],\end{aligned}\tag{C.4}$$

uniformly over $[0,1]$. Let

$$A=A_{\alpha,h}=\left[A_{1},\ldots A_{J_{L}}\right]=\mathbb{E}^{-1}\left[\frac{\mathbb{I}\left(I_{\ell}=I\right)\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}P\left(x_{\ell},t\right)P\left(x_{\ell},t\right)'K\left(t\right)dt}{B^{\left(1\right)}\left(\alpha|x_{\ell},I_{\ell}\right)}\right]$$

be a a $J_L\times J_L$ matrix with columns $A_j$, $j=1,\ldots,J_1$, $\left|A_j\right|_1$ the associated $\ell_1$ norm and $\left|A\right|_{1,\infty}=\max_{j\leq J_L}\left|A_j\right|_1$, $S$ a selection matrix which selects some columns of $A$, $a$, $b$ some conformable vectors and $|a|_{\infty}$ the largest entry of $a$.

$$|a'SAb| = \left|\sum_j b_j a' [SA]_j\right| \leq \sum_j |b_j| \max_j \left|a' [SA]_j\right| \leq |b|_1 |A|_{1,\infty} |a|_{\infty}.$$

This gives, since $\max_{\alpha,L} |A|_{1,\infty} < \infty$ by Lemma C.3 and by Assumption R-(ii),

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} |P'(x) S\mathsf{bias}_h(\alpha|I)| \\
&\qquad \leq C\left(\max_{x\in\mathcal{X}} \sum_{k=1}^{K} |P_k(x)|\right) \times \max_{1\leq k\leq K} \int |P_k(x)|\, dx = O(1), \\
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P'(x) SA\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} |P(x_\ell, t)| K(t)\, dt\right]\right| \\
&\qquad \leq C\left(\max_{x\in\mathcal{X}} \|P(x)\|\right) \times \max_{1\leq k\leq K} \int |P_k(x)|\, dx = O(1).
\end{aligned}$$

Let $S_0$ and $S_1$ be the selection matrices $S_0\mathsf{b} = \beta_0$ and $S_1\mathsf{b} = h\beta_1$, so that $\overline{B}(\alpha|x,I) = P'(x) S_0\overline{\mathsf{b}}(\alpha|I)$ and $\overline{B}^{(1)}(\alpha|x,I) = P'(x) S_1\overline{\mathsf{b}}(\alpha|I)/h$. Then (C.3), (C.4), Lemma B.1-(iii) and the above imply

$$\begin{aligned}
\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|\overline{B}(\alpha|x,I) - B(\alpha|x,I)\right| &\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|P'(x) S_0\left(\overline{\mathsf{b}}(\alpha|I) - \mathsf{b}^*(\alpha|I)\right)\right| \\
&\quad + \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} |\Psi(0|x, \mathsf{b}^*(\alpha|I)) - B(\alpha|x,I)| \\
&= o\left(h^{s+1}\right), \\
\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}} \left|\overline{B}^{(1)}(\alpha|x,I) - B^{(1)}(\alpha|x,I)\right| &= o(h^s),
\end{aligned}$$

$$
\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\alpha\left(\overline{B}^{(1)}(\alpha|x,I)-B^{(1)}(\alpha|x,I)\right)-h^{s+1}P'(x)\,\alpha S_1\mathsf{bias}_h(\alpha|I)\right|\\
&=\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\frac{1}{h}\left|\alpha P'(x)\,S_1\left(\overline{\mathsf{b}}(\alpha|I)-\mathsf{b}^*(\alpha|I)-h^{s+2}P'(x)\,\mathsf{bias}_h(\alpha|I)\right)\right|\\
&+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\frac{1}{h}\left|\alpha\left(P'(x)\,\mathsf{b}_1^*(\alpha|I)-hB^{(1)}(\alpha|x,I)\right)\right|\\
&=o\left(h^{s+1}\right).
\end{aligned}
$$

This ends the proof of the Theorem since $\overline{V}(\alpha|x,I)=\overline{B}(\alpha|x,I)+\alpha\overline{B}^{(1)}(\alpha|x,I)/(I-1)$.$\square$

# References


[1] Gragg, W.B. & R.A. Tapia (1974). Optimal error bounds for the Newton-Kantorovich Theorem. *SIAM Journal on Numerical Analysis* **11**, 10–13.

# Online Appendix D: Bahadur representation

Let $\widehat{\mathsf{e}}(\alpha|I)$ be a candidate linearization leading term for $\widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I)$ and $\widehat{\mathsf{d}}(\alpha|I)$ the associate linearization error term, or Bahadur remainder term,

$$\widehat{\mathsf{e}}(\alpha|I) = -\left(\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right)^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right), \tag{D.1}$$

$$\widehat{\mathsf{d}}(\alpha|I) = \widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I) - \widehat{\mathsf{e}}(\alpha|I). \tag{D.2}$$

This section goal is to study the magnitude of $\widehat{\mathsf{d}}(\alpha|I)$ and, in the ASQR case, the magnitude of $P'(x)\widehat{\mathsf{d}}_0(\alpha|I)$ and $P'(x)\widehat{\mathsf{d}}_1(\alpha|I)/h$.

**Theorem D.1** *Suppose Assumptions A, R-(i,ii) and S hold, $s \geq D_{\mathcal{M}}/2$ and*

$$\frac{\log L}{Lh^{2(D_{\mathcal{M}}+1)}} = o(1).$$

*Then*

$$\max_{\alpha\in[0,1]}\left\|\frac{Lh^{D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)/2}}{(h+\alpha(1-\alpha))^{1/2}\log L}\left\{\widehat{\mathsf{b}}(\alpha|I) - \overline{\mathsf{b}}(\alpha|I)\right.\right.$$
$$\left.\left. + \left(\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right)^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right\}\right\| = O_{\mathbb{P}}(1)$$

*with a diverging normalization term $Lh^{D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)/2}/\log L$. Moreover, for $\widehat{\mathsf{d}}(\alpha|I)$ as in (D.2),*

$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}\left\|P'(x)\widehat{\mathsf{d}}_0(\alpha|I)\right\| = O_{\mathbb{P}}\left(\frac{h^{1/2}\log L}{\left(Lh^{2D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)}\right)^{1/2}}\right),$$
$$\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}\left\|P'(x)\frac{\widehat{\mathsf{d}}_1(\alpha|I)}{h}\right\| = O_{\mathbb{P}}\left(\frac{\log L}{\left(Lh^{2D_{\mathcal{M}}+1+(D_{\mathcal{M}}\vee 1)}\right)^{1/2}}\right).$$

**Proof of Theorem D.1.** We first introduce some renormalizations. Let, for $\widehat{\mathsf{e}}(\alpha|I)$ as in (D.1),

$$\varrho_{\alpha L} = \frac{\left(h + \alpha\left(1-\alpha\right)\right)^{1/2} \log L}{L h^{D_{\mathcal{M}} + \left(D_{\mathcal{M}} \vee 1\right)/2}},$$

$$\widehat{\boldsymbol{R}}\left(d; \alpha, I\right) = \widehat{\mathsf{R}}\left(\overline{\mathsf{b}}\left(\alpha|I\right) + \widehat{\mathsf{e}}\left(\alpha|I\right) + \varrho_{\alpha L} d; \alpha, I\right) - \widehat{\mathsf{R}}\left(\overline{\mathsf{b}}\left(\alpha|I\right) + \widehat{\mathsf{e}}\left(\alpha|I\right); \alpha, I\right),$$

which is such that $\varrho_{\alpha L} = o\left(1\right)$ by $\log L / \left(L h^{2\left(D_{\mathcal{M}}+1\right)}\right) = o\left(1\right)$

$$\frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}} = \arg\min_{d} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right).$$

It follows that,

$$\begin{aligned}
\left\{\sup_{\alpha\in[0,1]} \left\| \frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}} \right\| \geq t\right\} &= \bigcup_{\alpha\in[0,1]} \left\{\left\| \frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}} \right\| \geq t\right\} \\
&\subset \bigcup_{\alpha\in[0,1]} \left\{\inf_{\|d\|\geq t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right) \leq \inf_{\|d\|\leq t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right)\right\} \subset \bigcup_{\alpha\in[0,1]} \left\{\inf_{\|d\|\geq t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right) \leq 0\right\}
\end{aligned}$$

since $\inf_{\|d\|\leq t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right) \leq \widehat{\boldsymbol{R}}\left(0; \alpha, I\right) = 0$. The next step uses a convexity argument that can be found in Pollard (1991). For any $d$ with $\|d\| \geq t$, convexity yields

$$\begin{aligned}
\widehat{\boldsymbol{R}}\left(d; \alpha, I\right) &= \frac{\|d\|}{t}\left\{\frac{t}{\|d\|}\widehat{\boldsymbol{R}}\left(\|d\| \frac{d}{\|d\|}; \alpha, I\right) + \left(1 - \frac{t}{\|d\|}\right)\widehat{\boldsymbol{R}}\left(0; \alpha, I\right)\right\} \\
&\geq \frac{\|d\|}{t}\widehat{\boldsymbol{R}}\left(t\frac{d}{\|d\|}; \alpha, I\right)
\end{aligned}$$

so that $\inf_{\|d\|\geq t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right) \leq 0$ implies $\inf_{\|d\|=t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right) \leq 0$ and them

$$\left\{\sup_{\alpha\in[0,1]} \left\| \frac{\widehat{\mathsf{d}}\left(\alpha|I\right)}{\varrho_{\alpha L}} \right\| \geq t\right\} \subset \left\{\inf_{\alpha\in[0,1]} \inf_{\|d\|=t} \widehat{\boldsymbol{R}}\left(d; \alpha, I\right) \leq 0\right\}. \tag{D.3}$$

Thus it is sufficient to consider those $d$ with $\|d\| = t$. The expression of $\widehat{\boldsymbol{R}}\left(d; \alpha, I\right)$ gives, using two Taylor expansions with integral remainder,

$$\begin{aligned}
\widehat{\boldsymbol{R}}(d;\alpha,I) &= \varrho_{\alpha L} d' \widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I);\alpha,I\right) \\
&\quad + \varrho_{\alpha L}^2 d' \left[\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha,I\right)(1-u)\,du\right] d' \\
&= \varrho_{\alpha L} d' \widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) \\
&\quad + \varrho_{\alpha L} d' \left[\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + u\widehat{\mathsf{e}}(\alpha|I);\alpha,I\right)du\right] \widehat{\mathsf{e}}(\alpha|I) \\
&\quad + \varrho_{\alpha L}^2 d' \left[\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha,I\right)(1-u)\,du\right] d'.
\end{aligned}$$

Since $\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) + \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\widehat{\mathsf{e}}(\alpha|I) = 0$ by (D.1), it follows that

$$\begin{aligned}
\widehat{\boldsymbol{R}}(d;\alpha,I) = \varrho_{\alpha L} d' &\left[\left\{\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + u\widehat{\mathsf{e}}(\alpha|I);\alpha,I\right) - \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right\}du\right] \widehat{\mathsf{e}}(\alpha|I) \\
&+ \varrho_{\alpha L}^2 d' \left[\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha,I\right)(1-u)\,du\right] d'.
\end{aligned}$$

Lemma B.4 and (C.3) with $s \geq D_{\mathcal{M}}/2$, $\log L / \left(L h^{2(D_{\mathcal{M}}+1)}\right) = o(1)$, Lemma B.2-(ii) give

$$\sup_{\alpha\in[0,1]} \left\| \frac{\widehat{\mathsf{e}}(\alpha|I)}{(h + \alpha(1-\alpha))^{1/2}} \right\| = O_{\mathbb{P}}\left(\left(\frac{\log L}{L h^{D_{\mathcal{M}}}}\right)^{1/2}\right) = o_{\mathbb{P}}\left(h^{D_{\mathcal{M}}/2+1}\right).$$

Lemmas B.3 and B.2-(i) then imply for the first item in $\widehat{\boldsymbol{R}}(d;\alpha,I)$, uniformly in $\alpha$ and $d$

with $\|d\| = t$,

$$\begin{aligned}
&\left|\varrho_{\alpha L} d' \left[\left\{\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + u\widehat{\mathsf{e}}(\alpha|I);\alpha, I\right) - \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right\} du\right] \widehat{\mathsf{e}}(\alpha|I)\right| \\
&\quad = \left|\varrho_{\alpha L} d' \left[\int_0^1 \left\{\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + u\widehat{\mathsf{e}}(\alpha|I);\alpha, I\right) - \overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right.\right.\right. \\
&\qquad\qquad\qquad \left.\left.\left. + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{d_{\mathcal{M}}+1}}\right)^{1/2}\right)\right\} du\right] \widehat{\mathsf{e}}(\alpha|I)\right| \\
&\quad = \left|\varrho_{\alpha L} d' \left[O_{\mathbb{P}}\left(h^{-D_{\mathcal{M}}/2}\right) \|\widehat{\mathsf{e}}(\alpha|I)\| + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right)\right] \widehat{\mathsf{e}}(\alpha|I)\right| \\
&\quad = t\left|\varrho_{\alpha L}\left[O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{2D_{\mathcal{M}}}}\right)^{1/2}\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right)\right] O_{\mathbb{P}}\left(\left(\frac{(h + \alpha(1-\alpha))\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right)\right| \\
&\quad = t\varrho_{\alpha L} O_{\mathbb{P}}\left(\frac{(h + \alpha(1-\alpha))^{1/2}\log L}{Lh^{D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)/2}}\right) = t\varrho_{\alpha L}^2 O_{\mathbb{P}}(1).
\end{aligned}$$

Observe that the condition $\log L / \left(Lh^{2(D_{\mathcal{M}}+1)}\right) = o(1)$ implies

$$\frac{\log L}{Lh^{D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)}} = o(1) \text{ and them } \varrho_{\alpha L} = o\left(\left(\frac{(h + \alpha(1-\alpha))\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right).$$

Lemmas B.3 and B.2 then imply for the second item in $\widehat{\boldsymbol{R}}(d;\alpha, I)$, uniformly in $\alpha$ and $d$ with $\|d\| = t$,

$$\begin{aligned}
&\varrho_{\alpha L}^2 d' \left[\int_0^1 \widehat{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha, I\right)(1-u)\, du\right] d' \\
&\quad = \varrho_{\alpha L}^2 d' \left[\int_0^1 \left\{\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I) + \widehat{\mathsf{e}}(\alpha|I) + u\varrho_{\alpha L} d;\alpha, I\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right)\right\}(1-u)\, du\right] d' \\
&\quad = \varrho_{\alpha L}^2 d' \left[\int_0^1 \left\{\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I)\right) + tO_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{2D_{\mathcal{M}}}}\right)^{1/2}\right) + O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right)\right\}(1-u)\, du\right] d' \\
&\quad \geq C\varrho_{\alpha L}^2 t^2 \left(1 + to_{\mathbb{P}}(1)\right).
\end{aligned}$$

Now (D.3) gives, with $O_{\mathbb{P}}(1)$ and $o_{\mathbb{P}}(1)$ which are uniform in $\alpha$,

$$\begin{aligned}\mathbb{P}\left(\sup_{\alpha\in[0,1]}\left\|\frac{\widehat{\mathsf{d}}(\alpha|I)}{\varrho_{\alpha L}}\right\|\geq t\right) &\leq \mathbb{P}\left(\inf_{\alpha\in[0,1]}\left\{C\varrho_{\alpha L}^2 t^2(1+to_{\mathbb{P}}(1))+t\varrho_{\alpha L}^2 O_{\mathbb{P}}(1)\right\}\leq 0\right)\\ &= \mathbb{P}(Ct(1+to_{\mathbb{P}}(1))+O_{\mathbb{P}}(1)\leq 0)\\ &\leq \mathbb{P}(t(1+to_{\mathbb{P}}(1))\leq |O_{\mathbb{P}}(1)|)\end{aligned}$$

which can be made as small as needed asymptotically by increasing $t$. This gives the first result of the Theorem. For the second and third, observe that $\max_{\alpha\in[0,1]}\varrho_{\alpha L}=\log L/Lh^{D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)/2}$ so that, uniformly in $\alpha$ and $x$,

$$\begin{aligned}\left|\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}P(x)'\widehat{\mathsf{d}}_0(\alpha|I)\right| &= (Lh)^{1/2}h^{D_{\mathcal{M}}/2}\max_{x\in\mathcal{X}}\|P(x)\|\left\|\widehat{\mathsf{d}}(\alpha|I)\right\|\\ &= O_{\mathbb{P}}\left((Lh)^{1/2}\varrho_{\alpha L}\right)=O_{\mathbb{P}}\left(\frac{h^{1/2}\log L}{\left(Lh^{2D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)}\right)^{1/2}}\right),\\ \left|\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}P(x)'\frac{\widehat{\mathsf{d}}_1(\alpha|I)}{h}\right| &= O_{\mathbb{P}}\left(\left(\frac{L}{h}\right)^{1/2}\varrho_{\alpha L}\right)=O_{\mathbb{P}}\left(\frac{\log L}{\left(Lh^{2D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)+1}\right)^{1/2}}\right).\end{aligned}$$

This ends the proof of the Theorem. □

# References

[1] Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7**, 186–199.

# Online Appendix E: Proof of main results

## E.1 Proof of Theorem 2

Recall that $s_1$ is the row vector $[0,1,0,\ldots,0]$ of dimension $s+2$ and let $s_0=[1,0,\ldots,0]$, $S_0=s_0\otimes\mathrm{Id}_K$, $S_1=s_1\otimes\mathrm{Id}_K$ so that $\widehat{\beta}_j(\alpha|I)=S_j\widehat{\beta}(\alpha|I)$, $j=0,1$ and

$$\widehat{V}(\alpha|x,I)=P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{b}}(\alpha|I),$$
$$\overline{V}(\alpha|x,I)=P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\overline{\mathsf{b}}(\alpha|I)$$

Define, for $\widehat{\mathsf{e}}(\alpha|I)$ as in (D.1)

$$\widehat{v}(\alpha|x,I)=\overline{V}(\alpha|x,I)+P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{e}}(\alpha|I) \tag{E.1}$$

which is such, for $\widehat{\mathsf{d}}(\alpha|I)$ as in (D.2),

$$\widehat{V}(\alpha|x,I)-\widehat{v}(\alpha|x,I)=P(x)'\left[S_0+\frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{d}}(\alpha|I).$$

As the eigenvalues of $\int_{\mathcal{X}}P(x)P(x)'dx$ are bounded away from infinity under Assumption R-(i)

$$\begin{aligned}\int_{\mathcal{X}}\int_0^1\left(\widehat{V}(\alpha|x,I)-\widehat{v}(\alpha|x,I)\right)^2d\alpha dx&=\frac{O\left(\sup_{\alpha\in[0,1]}\left\|\widehat{\mathsf{d}}(\alpha|I)\right\|^2\right)}{h^2}\\&=O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1+(D_{\mathcal{M}}\vee1)/2}}\right)^2\right)\end{aligned}$$

by Theorem D.1, which gives (4.5) since, by Assumption H,

$$\frac{Lh^{D_{\mathcal{M}}+1}}{\log L}\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1+(D_{\mathcal{M}}\vee1)/2}}\right)^2=\frac{\log L}{Lh^{D_{\mathcal{M}}+1+(D_{\mathcal{M}}\vee1)}}=o\left(\frac{\log L}{Lh^{2(D_{\mathcal{M}}+1)}}\right)=o(1).$$

That $\mathsf{bias}_{IL}^2 = O(1)$ and $\Sigma_{IL} = O(1)$ similarly follow from Assumption R-(i) and Proposition C.1-(i).

It holds since $\mathbb{E}[\widehat{\mathsf{e}}(\alpha|I)] = \overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)^{-1}\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right) = 0$ for all $\alpha$ in $[0,1]$

$$\mathbb{E}\left[\int_{\mathcal{X}}\int_0^1 \left(\widehat{v}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx\right] = \int_{\mathcal{X}}\int_0^1 \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx$$
$$+ \int_{\mathcal{X}}\int_0^1 \mathbb{E}\left[\left(P(x)'\left[S_0 + \frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{e}}(\alpha|I)\right)^2\right] d\alpha dx.$$

For the bias part, Theorem C.4 gives

$$\int_{\mathcal{X}}\int_0^1 \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx = \int_{\mathcal{X}}\int_0^1 \left(\frac{h^{s+1}P(x)'\alpha\mathsf{bias}_{1h}(\alpha|I)}{I-1} + o\left(h^{s+1}\right)\right)^2 d\alpha dx$$
$$= h^{2(s+1)}\int_{\mathcal{X}}\int_0^1 \left(\frac{P(x)'\alpha\mathsf{bias}_{1h}(\alpha|I)}{I-1}\right)^2 d\alpha dx + o\left(h^{2(s+1)}\right),$$

Since $\alpha\mathsf{bias}_{1h}(\alpha|I)/(I-1)$ differs from $\mathsf{bias}(\alpha|I)$ for $\alpha$ in $[0,h]$ or $[1-h,1]$, it follows

$$\int_{\mathcal{X}}\int_0^1 \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right)^2 d\alpha dx = h^{2(s+1)}\int_{\mathcal{X}}\int_0^1 \left(P(x)'\mathsf{bias}(\alpha|I)\right)^2 d\alpha dx + o\left(h^{2(s+1)}\right)$$
$$= h^{2(s+1)}\mathsf{bias}_{IL}^2 + o\left(h^{2(s+1)}\right).$$

Arguing similarly with Lemma B.5-(i) yields

$$\int_{\mathcal{X}}\int_0^1 \mathbb{E}\left[\left(P(x)'\left[S_0 + \frac{\alpha S_1}{h(I-1)}\right]\widehat{\mathsf{e}}(\alpha|I)\right)^2\right] d\alpha dx$$
$$= \int_{\mathcal{X}}\int_0^1 \mathbb{E}\left[\left(\left[\frac{P(x)'\alpha\widehat{\mathsf{e}}_1(\alpha|I)}{h(I-1)}\right]\right)^2\right] d\alpha dx + O\left(\frac{1}{Lh^{D_{\mathcal{M}}}}\right)$$
$$= \frac{\sigma_{LI}^2}{LIh^{D_{\mathcal{M}}+1}} + o\left(\frac{1}{Lh^{D_{\mathcal{M}}+1}}\right).$$

Substituting in the bias-variance decomposition of the integrated mean squared error ends the proof of the Theorem. □

## E.2 Proof of Theorem 3

Assumption R-(i) and Proposition C.1-(i) imply that $P(x)' \Sigma_h(\alpha|I) P(x) = 0$ holds only if $P(x) = 0$, which is impossible in the AQR case. But, in the ASQR case, if $P(x) = 0$ for some $x \in \mathcal{X}$ and all $K$ large enough, the approximation property S cannot hold, contradicting Assumption S-(ii). Assumptions R-(i), H and Proposition C.1-(i) imply

$$\max_{x \in \mathcal{X}} \left( P(x)' \Sigma_h(\alpha|I) P(x) \right) = O\left( \max_{x \in \mathcal{X}} \|P(x)\|^2 \right) = O\left( h^{-D_{\mathcal{M}}} \right).$$

By Theorem D.1, Lemma B.5, Assumptions R-(i), H, and using the same notations than in the proof of Theorem 2

$$\begin{aligned}
&\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2} \left( \widehat{V}(\alpha|x,I) - V(\alpha|x,I) - \frac{P'(x)\,\alpha S_1 \widehat{\mathsf{e}}(\alpha|I)}{h(I-1)} - \left(\overline{V}(\alpha|x,I) - V(\alpha|x,I)\right) \right) \\
&= \left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2} \left\{ P'(x)\,\widehat{\mathsf{e}}_0(\alpha|I) + P'(x) \left[ S_0 + \frac{\alpha S_1}{h(I-1)} \right] \widehat{\mathsf{d}}(\alpha|I) \right\} \\
&= \left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2} \left\{ O_{\mathbb{P}}\left( \frac{1}{\left(Lh^{D_{\mathcal{M}}}\right)^{1/2}} \right) + O\left( \frac{\left\| P(x)' \widehat{\mathsf{d}}(\alpha|I) \right\|}{h} \right) \right\} \\
&= O_{\mathbb{P}}\left( h^{1/2} + \left( \frac{\log^2 L}{Lh^{2D_{\mathcal{M}}-1+(D_{\mathcal{M}} \vee 1)}} \right)^{1/2} \right) = o_{\mathbb{P}}(1).
\end{aligned}$$

Since $\overline{V}(\alpha|x,I) - V(\alpha|x,I) = h^{s+1} P(x)' \mathsf{Bias}_h(\alpha|I) + o(h^{s+1})$, it remains to show that

$$\left( \frac{LIh}{P(x)' \Sigma_h(\alpha|I) P(x)} \right)^{1/2} \frac{\alpha P(x)' S_1 \widehat{\mathsf{e}}(\alpha|I)}{h(I-1)} \xrightarrow{d} \mathcal{N}(0,1).$$

Write

$$\left( \frac{LIh}{P(x)' \Sigma_h(\alpha|I) P(x)} \right)^{1/2} \frac{\alpha P(x)' S_1 \widehat{\mathsf{e}}(\alpha|I)}{h(I-1)} = \sum_{\ell=1}^{L} r_\ell(\alpha|x,I)$$

with $r_\ell(\alpha|x,I) = \mathbb{I}(I_\ell = I)\sum_{i=1}^{I_\ell} r_{i\ell}(\alpha|x,I)$ and

$$r_{i\ell}(\alpha|x,I) = \left(\frac{\alpha^2}{LIh(I-1)^2}\right)^{1/2} \frac{P(x)'}{\left(P(x)'\Sigma_h(\alpha|I)P(x)\right)^{1/2}} S_1\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|x,I);\alpha,I\right)\right]^{-1}$$
$$\times \int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}} \left\{\mathbb{I}\left(B_{i\ell} \leq P(x_\ell,t)\overline{\mathsf{b}}(\alpha|x,I)\right) - (\alpha+ht)\right\} P(x_\ell,t)K(t)\,dt.$$

Since $\mathbb{E}[r_\ell(\alpha|x,I)] = 0$ and $\max_{1\leq\ell\leq L}|\mathrm{Var}(r_\ell(\alpha|x,I)) - 1| = o(1)$, it is sufficient to show that $\max_{1\leq\ell\leq L}|\mathbb{E}[r_\ell^3(\alpha|x,I)]| = o(1)$ holds, see e.g. Theorem <19> p.179 in Pollard (2002). But Assumption R-(i) and Proposition C.1-(i), Lemma B.2 and (C.3),

$$|r_{i\ell}(\alpha|x,I)| \leq \frac{C}{(Lh)^{1/2}}\frac{\|P(x)\|}{\|P(x)\|} \times \max_{x\in\mathcal{X}}\|P(x)\| = O\left(\frac{1}{\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}}\right).$$

It follows that by Assumption H

$$\max_{1\leq\ell\leq L}\left|\mathbb{E}\left[r_\ell^3(\alpha|x,I)\right]\right| \leq I \max_{1\leq\ell\leq L, 1\leq i\leq I_\ell}|r_{i\ell}(\alpha|x,I)| \max_{1\leq\ell\leq L}\left|\mathbb{E}\left[r_\ell^2(\alpha|x,I)\right]\right|$$
$$= O\left(\frac{1}{\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}}\right) = o(1).$$

This ends the proof of the Theorem. □

## E.3 Proof of Theorem 4

The proof of Theorem requests some specific additional results. The next Lemma gives an expansion for

$$\mathcal{C}_h = \int_0^1\int_0^1 f(\alpha_1)g(\alpha_2)\left\{\int_0^1\int_0^1 \frac{1}{h}\pi\left(\frac{a_2-\alpha_2}{h}\right)K\left(\frac{a_2-\alpha_2}{h}\right)\right.$$
$$\left.\times\frac{1}{h}\pi'\left(\frac{a_1-\alpha_1}{h}\right)K\left(\frac{a_1-\alpha_1}{h}\right)[a_1\wedge a_2 - a_1a_2]\,da_1da_2\right\}d\alpha_1d\alpha_2.$$

Recall that $s_0' = [1,0,\ldots,0]$, $s_1' = [0,1,0,\ldots,0]$ and $s_2' = [0,0,1,0,\ldots,0]$ are vectors of dimension $s+2$.

**Lemma E.1** *Suppose that Assumption H holds. Assume that $f(\cdot) = f_h(\cdot)$ and $g(\cdot) = g_h(\cdot)$ are continuously differentiable functions, with, when $h$ goes to 0,*

$$\sup_{\alpha\in[0,1]} |f(\alpha)| = O(1) \ \ \textit{and} \ \sup_{\alpha\in[0,1]} |g(\alpha)| = O(1),$$
$$\sup_{\alpha\in[h,1-h]} \left|f^{(1)}(\alpha)\right| = O(1) \ \ \textit{and} \ \sup_{\alpha\in[h,1-h]} \left|g^{(1)}(\alpha)\right| = O(1),$$
$$\sup_{\alpha\in[0,h]\cup[1-h,1]} \left|f^{(1)}(\alpha)\right| = O\left(\frac{1}{h}\right) \ \ \textit{and} \ \sup_{\alpha\in[0,h]\cup[1-h,1]} \left|g^{(1)}(\alpha)\right| = O\left(\frac{1}{h}\right).$$

*Then, if $A$ is a random variable with a uniform distribution over* $[0,1]$

$$\begin{aligned}
\mathcal{C}_h =& \operatorname{Cov}\left(\left[\int_A^1 g(a)\,\Omega_h(a)\,da\right] s_0, \left[\int_A^1 f(a)\,\Omega_h(a)\,da\right] s_0\right) \\
&+ h\left\{\operatorname{Cov}\left(g(A)\,\Omega_h(A)\,s_1, \left[\int_A^1 f(a)\,\Omega_h(a)\,da\right] s_0\right)\right. \\
&\qquad \left. + \operatorname{Cov}\left(\left[\int_A^1 g(a)\,\Omega_h(a)\,da\right] s_0, f(A)\,\Omega_h(A)\,s_1\right)\right\} \\
&+ h^2 \operatorname{Cov}\left(g(A)\,\Omega_h(A)\,s_1, f(A)\,\Omega_h(A)\,s_1\right) \\
&- \frac{h^2}{2}\mathbb{E}\left[f(A)\,\Omega_h(A)\left[s_0 s_2' + s_2 s_0'\right] g(A)\,\Omega_h(A)\right] + o\left(h^2\right).
\end{aligned}$$

**Proof of Lemma E.1:** See Appendix F.

Consider two functions $\varphi_0(\alpha, x)$ and $\varphi_1(\alpha|x)$ and define

$$\begin{aligned}
\widehat{\boldsymbol{I}}_\varphi(x|I) = & \int_0^1 \left[\varphi_0(\alpha|x)\,s_0' + \varphi_1(\alpha|x)\,\frac{s_1'}{h}\right] \otimes P'(x) \\
& \times \left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right]^{-1} \widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right) d\alpha.
\end{aligned}$$

The purpose of the next Lemma is to compute the variance of this integral. Define for this purpose

$$\begin{aligned}
\mathbf{P} = \mathbf{P}(I) &= \mathbb{E}\left[\mathbb{I}(I_\ell = I) P(x_\ell) P(x_\ell)\right], \\
\mathbf{P}_0(\alpha) = \mathbf{P}_0(\alpha|I) &= \mathbb{E}\left[P(x_\ell) P(x_\ell) \frac{\mathbb{I}(I_\ell = I)}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right], \\
\mathbf{P}_1(\alpha) = \mathbf{P}_1(\alpha|I) &= -\mathbb{E}\left[P(x_\ell) P(x_\ell) \frac{\mathbb{I}(I_\ell = I) B^{(2)}(\alpha|x_\ell, I_\ell)}{\left(B^{(1)}(\alpha|x_\ell, I_\ell)\right)^2}\right],
\end{aligned}$$

and set

$$\mathsf{M}_0(\alpha) = \Omega_h(\alpha) \otimes \mathbf{P}_0(\alpha), \ \mathsf{M}_1(\alpha) = \Omega_{1h}(\alpha) \otimes \mathbf{P}_1(\alpha).$$

**Lemma E.2** *Suppose $s \geq D_{\mathcal{M}}/2$, and that Assumptions A, H, S and R hold. Assume that $\varphi_0(\alpha|x)$, $\varphi_1(\alpha|x)$ and $\frac{\partial \varphi_1(\alpha|x)}{\partial \alpha}$ are continuous functions in $(\alpha, x) \in [0,1] \times \mathcal{X}$. Let $A$ be a random variable with a uniform distribution over $[0,1]$. Then $\mathrm{Var}\left(\sqrt{LIh^{D_{\mathcal{M}}}}\widehat{\boldsymbol{I}}_\varphi(x|I)\right) = \sigma_L^2(x|I) + \left\|h^{D_{\mathcal{M}}/2} P(x)\right\|^2 o(1)$ with*

$$\sigma_L^2(x|I) = \mathrm{Var}\left[h^{D_{\mathcal{M}}/2} P'(x) \int_0^A \left(\varphi_0(\alpha|x) - \frac{\partial \varphi_1(\alpha|x)}{\partial \alpha}\right) \mathbf{P}_0(\alpha|I)^{-1} \mathbf{P}(I)^{1/2} d\alpha\right]$$

*and* $\mathrm{Var}\left(\sqrt{LI} \int_{\mathcal{X}} \widehat{\boldsymbol{I}}_\varphi(x|I)\, dx\right) = \sigma_L^2(I) + o(1)$ *with*

$$\sigma_L^2(I) = \mathrm{Var}\left[\int_0^A \left\{\int_{\mathcal{X}} P'(x) \left(\varphi_0(\alpha|x) - \frac{\partial \varphi_1(\alpha|x)}{\partial \alpha}\right) dx\right\} \mathbf{P}_0(\alpha|I)^{-1} \mathbf{P}^{1/2}(I)\, d\alpha\right].$$

**Proof of Lemma E.2.** Abbreviate $\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right)$, $\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I); \alpha, I\right)$ into $\overline{\mathsf{R}}^{(2)}(\alpha)$ and $\widehat{\mathsf{R}}^{(1)}(\alpha)$ respectively. We now give a suitable expansion for $\overline{R}^{(2)}(\alpha)^{-1}$. From the end of the proof of Lemma B.2 and Theorem C.4, it holds

$$\begin{aligned}
\overline{R}^{(2)}(\alpha) = \int \Bigg[\int_{\underline{I}_{\alpha,h} + o(h^s)}^{\overline{I}_{\alpha,h} + o(h^s)} \pi(t) \pi(t)' K(t)\, g\left[B(\alpha + ht|x, I) + o\left(h^{s+1}\right)|x, I\right] dt\Bigg] \\
\otimes P(x) P(x)' f(x, I)\, dx.
\end{aligned}$$

Since $s \geq 1$, $B^{(1)}(\cdot|x, I)$ is continuously differentiable. A first-order Taylor expansion gives that, uniformly,

$$\overline{R}^{(2)}(\alpha) = \mathsf{M}_0(\alpha) + h\mathsf{M}_1(\alpha) + o(h).$$

It then follows, uniformly over $[0,1]$

$$\begin{aligned}\left[\overline{R}^{(2)}(\alpha)\right]^{-1} &= \left[\mathrm{Id} + h\mathsf{M}_0(\alpha)^{-1}\mathsf{M}_1(\alpha) + o(h)\,\mathrm{Id}\right]^{-1}\mathsf{M}_0(\alpha)^{-1} \\ &= \mathsf{M}_0(\alpha)^{-1} - h\mathsf{M}_0(\alpha)^{-1}\mathsf{M}_1(\alpha)\,\mathsf{M}_0(\alpha)^{-1} + o(h)\,\mathrm{Id}.\end{aligned}$$

Now $\mathsf{M}_0(\alpha)^{-1} = \Omega_h(\alpha)^{-1} \otimes \mathbf{P}_0(\alpha)^{-1}$ and

$$\mathsf{M}_0(\alpha)^{-1}\mathsf{M}_1(\alpha)\,\mathsf{M}_0(\alpha)^{-1} = \left[\Omega_h(\alpha)^{-1}\Omega_{1h}(\alpha)\,\Omega_h(\alpha)^{-1}\right] \otimes \left[\mathbf{P}_0(\alpha)^{-1}\mathbf{P}_1(\alpha)\,\mathbf{P}_0(\alpha)^{-1}\right]$$

with

$$s_1'\Omega_h(\alpha)^{-1}\Omega_{1h}(\alpha) = s_1'\begin{bmatrix} 0 & 0 & \cdots & 0 & \times \\ 1 & 0 & & \vdots & c(\alpha) \\ 0 & 1 & & \vdots & \times \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & \times \end{bmatrix} = s_0' + c(\alpha)\,s_p$$

where $c(\alpha) = c_h(\alpha)$ and the entries of $\Omega_h(\alpha)^{-1}$ satisfy the smoothness conditions of Lemma E.1. This gives since the eigenvalues of $\Omega_h(\alpha)^{-1}$ and $\mathbf{P}_0(\alpha)^{-1}$ are bounded away from infinity uniformly in $\alpha$

$$\begin{aligned}\mathrm{Var}^{1/2}\left(\sqrt{LI}\widehat{\boldsymbol{I}}_\varphi(x|I)\right) &= \mathrm{Var}^{1/2}\left(\widehat{\boldsymbol{I}}_0(x|I) + \widehat{\boldsymbol{I}}_1(x|I) + \widehat{\boldsymbol{I}}_2(x|I) + \widehat{\boldsymbol{I}}_p(x|I)\right) \\ &\quad + o(1)\,\|P(x)\|\left\|\mathrm{Var}^{1/2}\left(\sqrt{LI}\int_0^1 \widehat{R}^{(1)}(\alpha)\,d\alpha\right)\right\|\end{aligned}$$

with

$$\widehat{\boldsymbol{I}}_0\left(x|I\right)=\sqrt{LI}\int_0^1\varphi_0\left(\alpha|x\right)\left[s_0\otimes P\left(x\right)\right]'\left[\Omega_h\left(\alpha\right)^{-1}\otimes\mathbf{P}_0\left(\alpha\right)^{-1}\right]\widehat{R}^{(1)}\left(\alpha\right)d\alpha,$$

$$\widehat{\boldsymbol{I}}_1\left(x|I\right)=-\sqrt{LI}\int_0^1\varphi_1\left(\alpha|x\right)\left[s_0\otimes P\left(x\right)\right]'\left[\Omega_h\left(\alpha\right)^{-1}\otimes\mathbf{P}_0\left(\alpha\right)^{-1}\mathbf{P}_1\left(\alpha\right)\mathbf{P}_0\left(\alpha\right)^{-1}\right]\widehat{R}^{(1)}\left(\alpha\right)d\alpha.$$

$$\widehat{\boldsymbol{I}}_2\left(x|I\right)=\sqrt{LI}\int_0^1\varphi_1\left(\alpha|x\right)\left[\frac{s_1}{h}\otimes P\left(x\right)\right]'\left[\Omega_h\left(\alpha\right)^{-1}\otimes\mathbf{P}_0\left(\alpha\right)^{-1}\right]\widehat{R}^{(1)}\left(\alpha\right)d\alpha,$$

$$\widehat{\boldsymbol{I}}_p\left(x|I\right)=\sqrt{LI}\int_0^1\varphi_1\left(\alpha|x\right)c\left(\alpha\right)\left[s_p\otimes P\left(x\right)\right]'\left[\Omega_h\left(\alpha\right)^{-1}\otimes\mathbf{P}_0\left(\alpha\right)^{-1}\mathbf{P}_1\left(\alpha\right)\mathbf{P}_0\left(\alpha\right)^{-1}\right]\widehat{R}^{(1)}\left(\alpha\right)d\alpha.$$

Observe now that, for any functions $f\left(\cdot\right)$ and $g\left(\cdot\right)$ satisfying the conditions of Lemma E.1

$$\begin{aligned}C_h\left(f,g\right)=\mathbb{E}\Bigg[&\mathbb{I}\left(I_\ell=I\right)\int_0^1\int_0^1 f\left(\alpha_1\right)g\left(\alpha_2\right)\\&\times\left\{G\left[\min\left(P\left(x_\ell,\frac{a_1-\alpha_1}{h}\right)\overline{\mathsf{b}}\left(\alpha_1|I\right),P\left(x_\ell,\frac{a_2-\alpha_2}{h}\right)\overline{\mathsf{b}}\left(\alpha_2|I\right)\right)\Big|\,x_\ell,I\right]\right.\\&\qquad\left.-G\left[P\left(x_\ell,\frac{a_1-\alpha_1}{h}\right)\overline{\mathsf{b}}\left(\alpha_1|I\right)\Big|\,x_\ell,I\right]G\left[P\left(x_\ell,\frac{a_2-\alpha_2}{h}\right)\overline{\mathsf{b}}\left(a_2|I\right)\Big|\,x_\ell,I\right]\right\}\\&\times\overline{R}^{(2)}\left(\alpha_1\right)^{-1}\left\{\left[\pi\left(\frac{a_1-\alpha_1}{h}\right)\pi\left(\frac{a_2-\alpha_2}{h}\right)'\right]\otimes\left[P\left(x_\ell\right)P\left(x_\ell\right)'\right]\right\}\overline{R}^{(2)}\left(\alpha_2\right)^{-1}\\&\times\frac{1}{h^2}K\left(\frac{a_1-\alpha_1}{h}\right)K\left(\frac{a_2-\alpha_2}{h}\right)'d\alpha_1 d\alpha_2\Bigg]\end{aligned}$$

Now (C.3), $\max_{(x,t)\in\mathcal{X}\times[-1,1]}\left\|P\left(x,t\right)\right\|=O\left(h^{-D_{\mathcal{M}}/2}\right)$ and Lemma B.1-(iii) gives

$$P\left(x_\ell,\frac{a-\alpha}{h}\right)\overline{\mathsf{b}}\left(\alpha|I\right)=B\left(a|x_\ell,I\right)+o\left(h^{s+1-D_{\mathcal{M}}/2}\right)$$

uniformly in $a$, $\alpha$ and $x_\ell$ with $\frac{a-\alpha}{h}$ in the support of $K\left(\cdot\right)$, $|a-\alpha|\leq h$. Since $s+1-D_{\mathcal{M}}/2\geq 0$, this gives under Assumption R-(ii) and by definition of $\mathbf{P}$

$$\begin{aligned}
C_h(f,g) = & \int_0^1 \int_0^1 f(\alpha_1) g(\alpha_2) \{a_1 \wedge a_2 - a_1 a_2\} \\
& \times \overline{R}^{(2)}(\alpha_1)^{-1} \left\{ \left[ \pi\left(\frac{a_1-\alpha_1}{h}\right) \pi\left(\frac{a_2-\alpha_2}{h}\right)' \right] \otimes \mathbf{P} \right\} \overline{R}^{(2)}(\alpha_2)^{-1} \\
& \times \frac{1}{h^2} K\left(\frac{a_1-\alpha_1}{h}\right) K\left(\frac{a_2-\alpha_2}{h}\right) d\alpha_1 d\alpha_2 dx_1 dx_2 \\
& + o(1)\,\mathrm{Id}\,.
\end{aligned}$$

Now applying Lemma E.1 gives, since $p \geq 2$

$$\operatorname{Var}\left(\widehat{\boldsymbol{I}}_p(x|I)\right) = \|P(x)\|^2 o(h), \quad \operatorname{Cov}\left(\widehat{\boldsymbol{I}}_p(x|I), \widehat{\boldsymbol{I}}_j(x|I)\right) = \|P(x)\|^2 o(1), \quad j = 1,2,3$$

$\left\|\operatorname{Var}\left(\sqrt{LI}\int_0^1 \widehat{R}^{(1)}(\alpha)\,d\alpha\right)\right\| = O(1)$ and

$$\begin{aligned}
& \operatorname{Var}\left(\widehat{\boldsymbol{I}}_0(x|I) + \widehat{\boldsymbol{I}}_1(x|I) + \widehat{\boldsymbol{I}}_2(x|I)\right) \\
& = P'(x) \left\{ \operatorname{Var}\left[ \int_0^A \left(\varphi_0(\alpha|x)\mathbf{P}_0(\alpha)^{-1} - \varphi_1(\alpha|x)\mathbf{P}_0(\alpha)^{-1}\mathbf{P}_1(\alpha)\mathbf{P}_0(\alpha)^{-1}\right) d\alpha \mathbf{P}^{1/2} \right] \right. \\
& \qquad - 2\operatorname{Cov}\left[ \int_0^A \left(\varphi_0(\alpha|x)\mathbf{P}_0(\alpha)^{-1} - \varphi_1(\alpha|x)\mathbf{P}_0(\alpha)^{-1}\mathbf{P}_1(\alpha)\mathbf{P}_0(\alpha)^{-1}\right) d\alpha \mathbf{P}^{1/2}, \right. \\
& \qquad\qquad \left. \left. \varphi_1(A|x)\mathbf{P}_0(A)^{-1}\mathbf{P}^{1/2} \right] + \operatorname{Var}\left[\varphi_1(A|x)\mathbf{P}_0(A)^{-1}\mathbf{P}^{1/2}\right] \right\} P(x) \\
& + o(1)\|P(x)\|^2 \\
& = \operatorname{Var}\left\{ P'(x) \left[ \int_0^A \left(\varphi_0(\alpha|x)\mathbf{P}_0(\alpha)^{-1} - \varphi_1(\alpha|x)\mathbf{P}_0(\alpha)^{-1}\mathbf{P}_1(\alpha)\mathbf{P}_0(\alpha)^{-1}\right) d\alpha \right.\right. \\
& \qquad \left.\left. -\varphi_1(A|x)\mathbf{P}_0(A)^{-1} \right] \mathbf{P}^{1/2} \right\} + o(1)\|P(x)\|^2 .
\end{aligned}$$

Observe now that

$$\frac{\partial}{\partial\alpha}\left[\varphi_1(\alpha|x)\mathbf{P}_0(\alpha)^{-1}\right] = \frac{\partial\varphi_1(\alpha|x)}{\partial\alpha}\mathbf{P}_0(\alpha)^{-1} - \varphi_1(\alpha|x)\mathbf{P}_0(\alpha)^{-1}\mathbf{P}_1(\alpha)\mathbf{P}_0(\alpha)^{-1}$$

so that

$$\int_0^A \left(\varphi_0(\alpha|x)\,\mathbf{P}_0(\alpha)^{-1} - \varphi_1(\alpha|x)\,\mathbf{P}_0(\alpha)^{-1}\mathbf{P}_1(\alpha)\,\mathbf{P}_0(\alpha)^{-1}\right)d\alpha - \varphi_1(A|x)\,\mathbf{P}_0(A)^{-1}$$
$$= \int_0^A \left(\varphi_0(\alpha|x) - \frac{\partial\varphi_1(\alpha|x)}{\partial\alpha}\right)\mathbf{P}_0(\alpha)^{-1}\,d\alpha + \varphi_1(0|x)\,\mathbf{P}_0(0)^{-1}.$$

This gives

$$\operatorname{Var}\left(\sqrt{LI}\widehat{\boldsymbol{I}}_{\varphi}(x|I)\right) = \operatorname{Var}\left\{P'(x)\int_0^A \left(\varphi_0(\alpha|x) - \frac{\partial\varphi_1(\alpha|x)}{\partial\alpha}\right)\mathbf{P}_0(\alpha)^{-1}\mathbf{P}^{1/2}d\alpha\right\} + o(1)\,\|P(x)\|^2$$

as stated in the first result of the Lemma. The second similarly follows, observing that $\left\|\int_{\mathcal{X}}\varphi_j(\alpha|x)\,P(x)\,dx\right\| = O(1)$, $j = 0, 1$under Assumption R-(ii). □

Consider two real valued continuous functions $\mathcal{F}_0(b_0, b_1)$ and $\mathcal{F}_1(b_0, b_1)$. Define

$$\varphi_0(\alpha|x,I) = \mathcal{F}_0\left(B(\alpha|x,I), B^{(1)}(\alpha|x,I)\right), \quad \varphi_1(\alpha|x,I) = \mathcal{F}_1\left(B(\alpha|x,I), B^{(1)}(\alpha|x,I)\right),$$
$$\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I) = \int_0^1 \left[\varphi_0(\alpha|x,I)\,s_0' + \varphi_1(\alpha|x,I)\,\frac{s_1'}{h}\right]\otimes P'(x)$$
$$\times\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)d\alpha.$$

A condition ensuring that the variances $\sigma_L^2(x|I)$ and $\sigma_L^2(I)$ of Lemma E.2 do not vanish is (4.9), that is

$$\varphi_0(\alpha|x,I) - \frac{\partial\varphi_1(\alpha|x,I)}{\partial\alpha} \neq 0.$$

**Proposition E.3** *Suppose $s \geq D_{\mathcal{M}}/2$, and that Assumptions A, H, S and R hold. Assume that $\varphi_0(\alpha|x)$, $\varphi_1(\alpha|x)$ and $\frac{\partial\varphi_1(\alpha|x)}{\partial\alpha}$ are continuous functions in $(\alpha, x) \in [0,1]\times\mathcal{X}$. Let $\sigma_L(x|I)$ and $\sigma_L(I)$ be as in Lemma E.2.*

*Then if (4.9) holds for some $\alpha$ of $[0,1]$ and if $Lh^{D_{\mathcal{M}}+2}$ diverges, $\sqrt{LIh^{D_{\mathcal{M}}}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)/\sigma_L(x|I)$ converges in distribution to a standard normal. If (4.9) holds for some $(\alpha, x)$ of $[0,1]\times\mathcal{X}$ and $Lh^2$ diverges, $\sqrt{LI}\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx/\sigma_L$ converges in distribution to a standard normal.*

**Proof of Proposition E.3.** The eigenvalues of $\mathbf{P}_0(\alpha)^{-1}$, $\mathbf{P}_1(\alpha)$ and $\mathbf{P}$ are bounded uniformly in $K$ and $\alpha$ by Assumptions R and S, and $\left\|h^{D_{\mathcal{M}}/2}P(x)\right\|$ is bounded away from 0 and infinity by Assumptions R and H. Then if (4.9) holds for some $\alpha$, $\sigma_L^2(x|I)$ is bounded away from 0 and infinity and the exact order of $\mathrm{Var}\left(\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\right)$ is $1/LIh^{D_{\mathcal{M}}}$. We now check the Lyapounov condition. Write $\widehat{\mathsf{R}}^{(1)}(\alpha)=\frac{1}{LI}\sum_{\ell=1}^{L}\mathbb{I}\left[I_\ell=I\right]r_\ell(\alpha)$, with

$$r_\ell(\alpha)=\sum_{i=1}^{I_\ell}\int_{-\frac{\alpha}{h}}^{\frac{1-\alpha}{h}}\left\{\mathbb{I}\left(B_{i\ell}\leq P(x_\ell,t)'\,\overline{\mathsf{b}}(\alpha|I)\right)-(\alpha+ht)\right\}\pi(t)\otimes P(x_\ell)K(t)\,dt.$$

This gives, since the eigenvalues of $\overline{\mathsf{R}}^{(2)}(\alpha)$ are asymptotically bounded from 0 by Lemma B.2 and (C.3),

$$\begin{aligned}&\mathbb{E}\left[\left|\int_0^1\left[\varphi_0(\alpha|x,I)\,s_0'+\varphi_1(\alpha|x,I)\,\frac{s_1'}{h}\right]\otimes P'(x)\left[\overline{\mathsf{R}}^{(2)}(\alpha)\right]^{-1}\frac{r_\ell(\alpha)-\mathbb{E}\left[r_\ell(\alpha)\right]}{LI}d\alpha\right|^3\right]\\&\quad\leq C\frac{h^{-1}\max_{x\in\mathcal{X}}\|P(x)\|^2}{(LI)^3}LI\,\mathrm{Var}\left(\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\right)=\frac{C}{L^2h^{D_{\mathcal{M}}+1}}\mathrm{Var}\left(\widehat{\boldsymbol{I}}(x|I)\right).\end{aligned}$$

$Lh^{D_{\mathcal{M}}+2}\to\infty$ implies that the Lyapounov condition holds since

$$\frac{C}{Lh^{D_{\mathcal{M}}+1}\,\mathrm{Var}^{3/2}\left(\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\right)}\mathrm{Var}\left(\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\right)=O\left(\frac{1}{\left(Lh^{D_{\mathcal{M}}+2}\right)^{1/2}}\right)\to 0$$

This implies that $\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)/\mathrm{Var}^{1/2}\left(\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\right)$ is asymptotically $\mathcal{N}(0,1)$, and them the stated asymptotic normality.

For $\sqrt{LI}\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx$, recall that $\left\|\int|P(x)|\,dx\right\|=O(1)$ by Assumption R. This also gives

$$\begin{aligned}&\mathbb{E}\left[\left|\int_{\mathcal{X}}\left[\int_0^1\left(\varphi_0(\alpha|x,I)\,s_0'+\varphi_1(\alpha|x,I)\,\frac{s_1'}{h}\right)\otimes P'(x)\right]\left[\overline{\mathsf{R}}^{(2)}(\alpha)\right]^{-1}\frac{r_\ell(\alpha)-\mathbb{E}\left[r_\ell(\alpha)\right]}{LI}d\alpha\right|^3\right]\\&\quad\leq C\frac{h^{-1}}{(LI)^3}LI\,\mathrm{Var}\left(\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx\right)=\frac{C}{L^2h}\mathrm{Var}\left(\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx\right).\end{aligned}$$

Therefore the Lyapounov condition holds since $Lh^2$ diverges, because

$$\frac{C}{Lh\,\mathrm{Var}^{3/2}\left(\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}\left(x|I\right)dx\right)}\mathrm{Var}\left(\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}\left(x|I\right)dx\right)=\frac{C}{\left(Lh^2\right)^{1/2}}\to 0$$

The rest of the proof is as above. □

**Proof of Theorem 4.** Let $\widehat{\mathsf{d}}\left(\alpha|I\right)$ and $\widehat{\mathsf{e}}\left(\alpha|I\right)$ be as in (D.2) and (D.1),

$$\begin{aligned}\widehat{\mathsf{e}}\left(\alpha|I\right)&=-\left(\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right)\right)^{-1}\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}\left(\alpha|I\right);\alpha,I\right),\\ \widehat{\mathsf{d}}\left(\alpha|I\right)&=\widehat{\mathsf{b}}\left(\alpha|I\right)-\overline{\mathsf{b}}\left(\alpha|I\right)-\widehat{\mathsf{e}}\left(\alpha|I\right).\end{aligned}$$

Let $\widehat{\boldsymbol{I}}_{\mathcal{F}}\left(x|I\right)$ be as above, replacing $\varphi_j\left(\cdot\right)$ with $\varphi_{jI}\left(\cdot\right)$, $j=0,1$. Then the second-order Taylor inequality gives

$$\begin{aligned}&\widehat{\theta}\left(x\right)-\theta\left(x\right)\\ &\quad=\sum_{I\in\mathcal{I}}\int_0^1\left[\varphi_{0I}\left(\alpha,x\right)\left(\overline{B}\left(\alpha|x,I\right)-B\left(\alpha|x,I\right)\right)+\varphi_{1I}\left(\alpha,x\right)\left(\overline{B}^{(1)}\left(\alpha|x,I\right)-B^{(1)}\left(\alpha|x,I\right)\right)\right]d\alpha\\ &\quad+\sum_{I\in\mathcal{I}}\widehat{\boldsymbol{I}}_{\mathcal{F}}\left(x|I\right)\\ &\quad+\sum_{I\in\mathcal{I}}\int_0^1\left[\left(\varphi_{0I}\left(\alpha,x\right)s_0'+\varphi_{1I}\left(\alpha,x\right)\frac{s_1'}{h}\right)\otimes P'\left(x\right)\right]\widehat{\mathsf{d}}\left(\alpha|I\right)d\alpha\\ &\quad+O\left(1\right)\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left[\left(\overline{B}\left(\alpha|x,I\right)-B\left(\alpha|x,I\right)\right)^2+\left(\overline{B}^{(1)}\left(\alpha|x,I\right)-B^{(1)}\left(\alpha|x,I\right)\right)^2\right]\\ &\quad O\left(1\right)\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left[\left(\left[s_0'\otimes P'\left(x\right)\right]\widehat{\mathsf{e}}\left(\alpha|I\right)\right)^2+\left(\left[\frac{s_1'}{h}\otimes P'\left(x\right)\right]\widehat{\mathsf{e}}\left(\alpha|I\right)\right)^2\right]\\ &\quad O\left(1\right)\sup_{(\alpha,x,I)\in[0,1]\times\mathcal{X}\times\mathcal{I}}\left[\left(\left[s_0'\otimes P'\left(x\right)\right]\widehat{\mathsf{d}}\left(\alpha|I\right)\right)^2+\left(\left[\frac{s_1'}{h}\otimes P'\left(x\right)\right]\widehat{\mathsf{d}}\left(\alpha|I\right)\right)^2\right].\end{aligned}$$

Theorems C.4 and D.1, Lemma B.5 give

$$\begin{aligned}
\widehat{\theta}(x)-\theta(x) &= o(h^s)+\sum_{I\in\mathcal{I}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I) \\
&\quad+\frac{1}{(Lh^{D_{\mathcal{M}}})^{1/2}}O_{\mathbb{P}}\left(\frac{\log L}{\left(Lh^{2D_{\mathcal{M}}+2+(D_{\mathcal{M}}\vee 1)}\right)^{1/2}}+\frac{\log L}{(Lh^{D_{\mathcal{M}}+2})^{1/2}}\right) \\
&= o(h^s)+\sum_{I\in\mathcal{I}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)+o_{\mathbb{P}}\left(\frac{1}{(Lh^{D_{\mathcal{M}}})^{1/2}}\right).
\end{aligned}$$

Proposition E.3 then gives the result since the $\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)$ are independent. The asymptotic normality of $\widehat{\theta}$ similarly follows from Assumption R, which gives $\left\|\int_{\mathcal{X}}|P(x)|\,dx\right\| = O(1)$, and Theorem D.1 which implies

$$\begin{aligned}
\widehat{\theta}-\theta &= o(h^s)+\sum_{I\in\mathcal{I}}\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx \\
&\quad+O\left(\frac{\sup_{\alpha\in[0,1]}\left\|\widehat{\mathsf{d}}(\alpha|I)\right\|}{h}\right)+\frac{1}{L^{1/2}}O_{\mathbb{P}}\left(\frac{\log L}{(Lh^{2D_{\mathcal{M}}+2})^{1/2}}\right) \\
&= o(h^s)+\sum_{I\in\mathcal{I}}\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx+\frac{1}{L^{1/2}}O_{\mathbb{P}}\left(\frac{\log L}{(Lh^{2D_{\mathcal{M}}+2})^{1/2}}\right) \\
&= o(h^s)+\sum_{I\in\mathcal{I}}\int_{\mathcal{X}}\widehat{\boldsymbol{I}}_{\mathcal{F}}(x|I)\,dx+o_{\mathbb{P}}\left(\frac{1}{L^{1/2}}\right). \qquad \square
\end{aligned}$$

## E.4 Proof of Theorem A.1

By Theorems C.4 and D.1, Lemma B.5 and using the notations of the proof of Theorem 2

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\widehat{B}\left(\alpha|x,I\right)-B\left(\alpha|x,I\right)\right| \\
&\quad\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'S_0\left[\widehat{\mathsf{b}}\left(\alpha|I\right)-\overline{\mathsf{b}}\left(\alpha|I\right)\right]\right|+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\overline{B}\left(\alpha|x,I\right)-B\left(\alpha|x,I\right)\right| \\
&\quad\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\widehat{\mathsf{e}}_0\left(\alpha|I\right)\right|+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left\|P\left(x\right)'\widehat{\mathsf{d}}_0\left(\alpha|I\right)\right\|+o\left(h^{s+1}\right) \\
&\quad= O_{\mathbb{P}}\left[\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\left\{1+\left(\frac{\log L}{Lh^{2D_{\mathcal{M}}+(D_{\mathcal{M}}\vee 1)}}\right)^{1/2}\right\}\right]+o\left(h^{s+1}\right) \\
&\quad= O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}\right)+o\left(h^{s+1}\right)
\end{aligned}$$

$$\begin{aligned}
&\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\widehat{V}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right| \\
&\quad\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\left(S_0+\frac{\alpha}{h}S_1\right)\left[\widehat{\mathsf{b}}\left(\alpha|I\right)-\overline{\mathsf{b}}\left(\alpha|I\right)\right]\right|+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\overline{V}\left(\alpha|x,I\right)-V\left(\alpha|x,I\right)\right| \\
&\quad\leq \sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\widehat{\mathsf{e}}_0\left(\alpha|I\right)\right|+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P\left(x\right)'\frac{\widehat{\mathsf{e}}_1\left(\alpha|I\right)}{h}\right| \\
&\qquad+\sup_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left\|P\left(x\right)'\left(\widehat{\mathsf{d}}_0+\alpha\frac{\widehat{\mathsf{d}}_1\left(\alpha|I\right)}{h}\right)\right\|+O\left(h^{s+1}\right) \\
&\quad= O_{\mathbb{P}}\left[\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\left\{1+\left(\frac{\log L}{Lh^{2D_{\mathcal{M}}+1+(D_{\mathcal{M}}\vee 1)}}\right)^{1/2}\right\}\right]+O\left(h^{s+1}\right) \\
&\quad= O_{\mathbb{P}}\left(\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}\right)+O\left(h^{s+1}\right).
\end{aligned}$$

This end the proof of the Theorem. □

# Online Appendix F: Proofs of intermediary results

## F.1 Lemmas B.1, B.2 and C.3

**Proof of Lemma B.1.** Consider the harder $ASQR$ case. (i) It holds that, for $\beta_k(\cdot|\cdot)$ as in (2.11),

$$\begin{aligned}
&B(\alpha+ht|x,I) - P(x,t)' \mathsf{b}^*(\alpha|I) \\
&= B(\alpha+ht|x,I) - \sum_{k=1}^{K} P_k(x)\beta_k(\alpha+ht|I) \\
&+ \sum_{k=1}^{K} P_k(x)\beta_k(\alpha+ht|I) - \sum_{k=1}^{K} P_k(x)\sum_{p=0}^{s+1}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I) \\
&= B(\alpha+ht|x,I) - \sum_{k=1}^{K} P_k(x)\beta_k(\alpha+ht|I) \\
&+ \sum_{k=1}^{K} P_k(x)\left(\beta_k(\alpha+ht|I) - \sum_{p=0}^{s}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I)\right) - \frac{(ht)^{s+1}}{(s+1)!}\sum_{k=1}^{K} P_k(x)\beta_k^{(s+1)}(\alpha|I).
\end{aligned}$$

A Taylor expansion with integral remainder gives

$$\beta_k(\alpha+ht|I) - \sum_{p=0}^{s}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I) = \frac{(ht)^{s+1}}{s!}\int_0^1 \beta_k^{(s+1)}(\alpha+uht|I)(1-u)^s\,du$$

so that

$$
\begin{aligned}
&B(\alpha+ht|x,I)-P(x,t)'\,\mathsf{b}^*(\alpha|I)\\
&\quad=B(\alpha+ht|x,I)-\sum_{k=1}^{K}P_k(x)\,\beta_k(\alpha+ht|I)\\
&\quad+\frac{(ht)^{s+1}}{s!}\int_0^1\left\{\sum_{k=1}^{K}P_k(x)\,\beta_k^{(s+1)}(\alpha+uht|I)-B^{(s+1)}(\alpha+uht|I)\right\}(1-u)^s\,du\\
&\quad+\frac{(ht)^{s+1}}{s!}\int_0^1\left\{B^{(s+1)}(\alpha+uht|x,I)-B^{(s+1)}(\alpha|x,I)\right\}(1-u)^s\,du\\
&\quad+\frac{(ht)^{s+1}}{(s+1)!}\left\{B^{(s+1)}(\alpha|x,I)-\sum_{k=1}^{K}P_k(x)\,\beta_k^{(s+1)}(\alpha|x,I)\right\}.
\end{aligned}
$$

Hence since $B^{(s+1)}(\alpha|x,I)$ is continuous, by Property S and Proposition C.1

$$
\begin{aligned}
\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|B(\alpha+ht|x,I)-P(x,t)\,\mathsf{b}^*(\alpha|I)\right|&=o\left(h^{s+1}\right)+o\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right)\\
&=o\left(h^{s+1}\right) \qquad \text{(F.1)}
\end{aligned}
$$

since $K^{-1/D_{\mathcal{M}}}=O(h)$. Observe also that, uniformly in $\alpha$, $x$ and $t$ as above,

$$
\begin{aligned}
\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}^*(\alpha|I)\right]&=\sum_{p=1}^{s+1}h^p\frac{t^{p-1}}{(p-1)!}\sum_{k=1}^{K}P_k(x)\,\beta_k^{(p)}(\alpha|I)\\
&=h\left(B^{(1)}(\alpha|x,I)+o(1)\right)+h^2\left(\sum_{p=2}^{s+1}h^{p-2}\frac{t^{p-1}}{(p-1)!}B^{(p)}(\alpha|x,I)+o(1)\right)\\
&=hB^{(1)}(\alpha|x,I)+o(h)
\end{aligned}
$$

by Property S, which also gives,

$$
\begin{aligned}
\max_{p=1,\ldots,s+1}\left(\frac{\max_{x\in\mathcal{X}}\left|P(x)'\,\mathsf{b}_p^*(\alpha|I)\right|}{h}\right)&=\max_{p=1,\ldots,s+1}\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}h^{p-1}\left|B^{(p)}(\alpha|x,I)+o(1)\right|\\
&=\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}B^{(1)}(\alpha|x,I)+o(1)\leq\overline{f}
\end{aligned}
$$

provided $\overline{f}$ is large enough and $h$ small enough, so that $\mathsf{b}^*(\alpha|I)$ is in $\underline{\mathcal{BI}}_{\alpha,h}$ since $B^{(1)}(\cdot|\cdot,\cdot)$ is bounded away from 0 and infinity by Proposition C.1. Suppose now that $\|\mathsf{b}-\mathsf{b}^*(\alpha|I)\| \leq Ch/K^{1/2} = Ch^{D_{\mathcal{M}}/2+1}$. Then

$$
\begin{aligned}
\left|\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}\right]\right| &\geq \left|\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}^*(\alpha|I)\right]\right| - \|\mathsf{b}-\mathsf{b}^*(\alpha|I)\|\,\|P(x)\| \\
&\geq \left|\frac{\partial}{\partial t}\left[P(x,t)'\,\mathsf{b}^*(\alpha|I)\right]\right| - O(h), \\
\left|P(x)'\,\mathsf{b}_p\right| &\leq \left|P(x)'\,\mathsf{b}_p^*(\alpha|I)\right| + \|\mathsf{b}-\mathsf{b}^*(\alpha|I)\|\,\|P(x)\| \\
&\leq \left|P(x)'\,\mathsf{b}_p^*(\alpha|I)\right| - Ch, \quad p=1,\ldots,s+1,
\end{aligned}
$$

and $\mathcal{B}\left(\mathsf{b}^*(\alpha|I), Ch^{D_{\mathcal{M}}/2+1}\right) \subset \underline{\mathcal{BI}}_{\alpha,h}$ when $h$ is small enough provided $C$ is small enough. Hence (i) holds. (ii) follows from the Implicit Function Theorem and the definition of $\mathcal{BI}_{\alpha,h}$.

The first equality of (iii) is (F.1). For the second, note that $\alpha + ht \geq h > 0$ when $\alpha \geq 3h$ for all $t$ in $\mathcal{I}_{\alpha,h}$. It holds

$$
\begin{aligned}
&B(\alpha+ht|x,I) - P(x,t)'\,\mathsf{b}^*(\alpha|I) \\
&\quad = B(\alpha+ht|x,I) - \sum_{k=1}^{K} P_k(x)\,\beta_k(\alpha+ht|I) \\
&\quad + \sum_{k=1}^{K} P_k(x)\left(\beta_k(\alpha+ht|I) - \sum_{p=0}^{s+1}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I)\right)
\end{aligned}
$$

with

$$
\beta_k(\alpha+ht|I) - \sum_{p=0}^{s+1}\frac{(ht)^p}{p!}\beta_k^{(p)}(\alpha|I) = \frac{(ht)^{s+2}}{(s+1)!}\int_0^1 \beta_k^{(s+2)}(\alpha+uht|I)\,(1-u)^{s+1}\,du
$$

recalling, as established in the proof of Proposition C.1-(i) for $\alpha > 0$,

$$
\begin{aligned}
\beta_k^{(s+2)}(\alpha|I) &= \frac{1}{\alpha}\Big((I-1)\,\gamma_k^{(s+1)}(\alpha|I) - (I+s)\,\beta_k^{(s+1)}(\alpha|I)\Big), \\
B^{(s+2)}(\alpha|x,I) &= \frac{1}{\alpha}\Big((I-1)\,V_k^{(s+1)}(\alpha|I) - (I+s)\,B^{(s+1)}(\alpha|x,I)\Big). \qquad \text{(F.2)}
\end{aligned}
$$

Hence

$$
\begin{aligned}
& B\left(\alpha+ht|x,I\right)-P\left(x,t\right)'\mathsf{b}^{*}\left(\alpha|I\right)-\frac{\left(ht\right)^{s+2}}{\left(s+2\right)!}B^{\left(s+2\right)}\left(\alpha|I\right) \\
& \quad = B\left(\alpha+ht|x,I\right)-\sum_{k=1}^{K}P_{k}\left(x\right)\beta_{k}\left(\alpha+ht|I\right) \\
& \quad + \frac{\left(ht\right)^{s+2}}{\left(s+1\right)!}\int_{0}^{1}\left\{\sum_{k=1}^{K}P_{k}\left(x\right)\beta_{k}^{\left(s+2\right)}\left(\alpha+uht|I\right)-B^{\left(s+2\right)}\left(\alpha+uht|x,I\right)\right\}\left(1-u\right)^{s+1}du \\
& \quad + \frac{\left(ht\right)^{s+2}}{\left(s+1\right)!}\int_{0}^{1}\left\{B^{\left(s+2\right)}\left(\alpha+uht|x,I\right)-B^{\left(s+2\right)}\left(\alpha|x,I\right)\right\}\left(1-u\right)^{s+1}du,
\end{aligned}
$$

with, using the expressions $\beta_{k}^{(s+2)}\left(\cdot|\cdot\right)$ and $B^{(s+2)}\left(\cdot|\cdot\right)$ of the proof of Proposition C.1

$$
\max_{(\alpha,x)\in[0,3h]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B\left(\alpha+ht|x,I\right)-\sum_{k=1}^{K}P_{k}\left(x\right)\beta_{k}\left(\alpha+ht|I\right)\right)\right| = ho\left(K^{-\frac{s+1}{D_{\mathcal{M}}}}\right)=o\left(h^{s+2}\right),
$$

$$
\begin{aligned}
& \max_{(\alpha,x)\in[3h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\int_{0}^{1}\left\{\sum_{k=1}^{K}P_{k}\left(x\right)\beta_{k}^{\left(s+2\right)}\left(\alpha+uht|I\right)-B^{\left(s+2\right)}\left(\alpha+uht|x,I\right)\right\}\left(1-u\right)^{s+1}du\right| \\
& \quad \leq C\max_{(\alpha,x)\in[2h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left\{\frac{\alpha}{\alpha-h}\left|\sum_{k=1}^{K}P_{k}\left(x\right)\beta_{k}^{\left(s+1\right)}\left(\alpha|I\right)-B\left(\alpha|x,I\right)\right|\right\} \\
& \quad + C\max_{(\alpha,x)\in[2h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left\{\frac{\alpha}{\alpha-h}\left|\sum_{k=1}^{K}P_{k}\left(x\right)\gamma_{k}^{\left(s+1\right)}\left(\alpha|I\right)-V\left(\alpha|x,I\right)\right|\right\}=o\left(1\right), \\
& \max_{(\alpha,x)\in[3h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\int_{0}^{1}\left\{B^{\left(s+2\right)}\left(\alpha+uht|x,I\right)-B^{\left(s+2\right)}\left(\alpha|x,I\right)\right\}\left(1-u\right)^{s+1}du\right|=o(1).
\end{aligned}
$$

Substituting gives

$$
\max_{(\alpha,x)\in[3h,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B\left(\alpha+ht|x,I\right)-P\left(x,t\right)'\mathsf{b}^{*}\left(\alpha|I\right)-\frac{\left(ht\right)^{s+2}}{\left(s+2\right)!}B^{\left(s+2\right)}\left(\alpha|x,I\right)\right)\right|=o\left(h^{s+2}\right)
$$

which implies the second statement in (iii) since by Proposition C.1-(i) and (C.3)

$$\max_{(\alpha,x)\in[0,3h]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\left(B\left(\alpha+ht|x,I\right)-P\left(x,t\right)'\mathsf{b}^*\left(\alpha|I\right)\right)\right|=o\left(h^{s+2}\right),$$

$$\max_{(\alpha,x)\in[0,3h]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\alpha\frac{(ht)^{s+2}}{(s+2)!}B^{(s+2)}\left(\alpha|x,I\right)\right|=o\left(h^{s+2}\right).$$

The third result in (iii) follows from Proposition C.1-(iii). The fourth equality of (iii) follows from

$$\begin{aligned}
o\left(h^{s+1}\right) &= \max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{t\in\mathcal{I}_{\alpha,h}}\left|\Psi\left(t|x,\mathsf{b}^*\left(\alpha|I\right)\right)-B\left(\alpha+ht|x,I\right)\right| \\
&= \max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|\Psi\left[\Delta\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)|x,\mathsf{b}^*\left(\alpha|I\right)\right]\right. \\
&\qquad\qquad \left.-B\left[\alpha+h\Delta\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)|x,I\right]\right| \\
&= \max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|u-B\left[\alpha+h\Delta\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)|x,I\right]\right| \\
&= \max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|B\left[\alpha+h\frac{G\left(u|x,I\right)-\alpha}{h}|x,I\right]\right. \\
&\qquad\qquad \left.-B\left[\alpha+h\Delta\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)|x,I\right]\right| \\
&\geq Ch\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\max_{u\in\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}^*(\alpha|I)\right]}\left|\frac{G\left(u|x,I\right)-\alpha}{h}-\frac{\Phi\left(u|x,\mathsf{b}^*\left(\alpha|I\right)\right)-\alpha}{h}\right|
\end{aligned}$$

by Proposition C.1-(i).

Consider now (iv). The first bound follows from the Cauchy-Schwarz inequality. This bound implies for all $u$ in $\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]\cap\Psi\left[\mathcal{I}_{\alpha,h}|x,\mathsf{b}_1\right]$

$$\begin{aligned}
&\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-\Psi\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right| \\
&\quad=\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-u\right| \\
&\quad=\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]\right|\leq Ch^{-D_{\mathcal{M}}/2}\left\|\mathsf{b}_1-\mathsf{b}_0\right\|.
\end{aligned}$$

By definition of $\underline{\mathcal{BI}}_{\alpha,h}$

$$
\begin{aligned}
&\left|\Psi\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_0\right]-\Psi\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right| \\
&\quad\geq Ch\left|\Delta\left(u|x,\mathsf{b}_1\right)-\Delta\left(u|x,\mathsf{b}_0\right)\right|=C\left|\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)\right|
\end{aligned}
$$

and substituting shows that the second bound of (iv) holds. For the third bound in (iv), it holds uniformly in $\alpha$, $x$, $u$, $\mathsf{b}_1$ and $\mathsf{b}_0$

$$
\begin{aligned}
&\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right| \\
&\quad\leq\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_1\right]\right| \\
&\quad+\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right| \\
&\quad\leq\max_{t\in\mathcal{I}_{\alpha,h}}\left|\frac{\partial^2\Psi\left(t|x,\mathsf{b}_1\right)}{\partial t^2}\right|\frac{\left|\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)\right|}{h} \\
&\quad+\max_{t\in\mathcal{I}_{\alpha,h}}\left|\frac{\partial P\left(x,t\right)}{\partial t}\left(\mathsf{b}_1-\mathsf{b}_0\right)\right|.
\end{aligned}
$$

But, by definition of $\underline{\mathcal{BI}}_{\alpha,h}$

$$
\max_{t\in\mathcal{I}_{\alpha,h}}\left|\frac{\partial^2\Psi\left(t|x,\mathsf{b}_1\right)}{\partial t^2}\right|\leq Ch\max_{p=2,\ldots,s+1}\left|\frac{P\left(x\right)\mathsf{b}_{1p}}{h}\right|=O\left(h\right)
$$

so that substituting and the bound for $\Phi\left(u|x,\mathsf{b}_1\right)-\Phi\left(u|x,\mathsf{b}_0\right)$ gives, uniformly in $\alpha$, $x$, $u$, $\mathsf{b}_1$ and $\mathsf{b}_0$

$$
\left|\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_1\right)|x,\mathsf{b}_1\right]-\frac{\partial\Psi}{\partial t}\left[\Delta\left(u|x,\mathsf{b}_0\right)|x,\mathsf{b}_0\right]\right|\leq Ch^{-D_{\mathcal{M}}/2}\left\|\mathsf{b}_1-\mathsf{b}_0\right\|,
$$

which is the fourth inequality. The expression in (ii) of $\Phi\left(\cdot\right)$ and the definition of $\underline{\mathcal{BI}}_{\alpha,h}$ yield the third inequality. $\square$

**Proof of Lemma B.2.** It holds

$$
\begin{aligned}
\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) &= \mathbb{E}\Bigg[\mathbb{I}\left[B_{i\ell}\in\Psi\left(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}\right),I_\ell=I\right] \\
&\qquad \frac{P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)P\left(x_\ell,\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)'}{\Psi\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)}K\left(\Delta\left(B_{i\ell}|x_\ell,\mathsf{b}\right)\right)\Bigg] \\
&= \int\left[\int_{\Psi\left(\underline{I}_{\alpha,h}|x,\mathsf{b}\right)\vee B(0|x,I)}^{\Psi\left(\overline{I}_{\alpha,h}|x,\mathsf{b}\right)\wedge B(1|x,I)}\frac{P\left(x,\Delta\left(y|x,\mathsf{b}\right)\right)P\left(x,\Delta\left(y|x,\mathsf{b}\right)\right)'}{\Psi\left(\Delta\left(y|x,\mathsf{b}\right)|x_\ell,\mathsf{b}\right)}K\left(\Delta\left(y|x,\mathsf{b}\right)\right)g\left(y,x,I\right)dy\right]dx.
\end{aligned}
$$

Recall $\Delta\left[\Psi\left[t|x,\mathsf{b}\right]|x,\mathsf{b}\right]=t$ for all $t$ in $\mathcal{I}_{\alpha,h}$ and let

$$
\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)=\overline{I}_{\alpha,h}\wedge\Delta\left[B\left(1|x,I\right)|x,\mathsf{b}\right],\quad \underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)=\underline{I}_{\alpha,h}\vee\Delta\left[B\left(0|x,I\right)|x,\mathsf{b}\right].
$$

The change of variable $y=\Psi\left(t|x,\mathsf{b}\right)$ yields that

$$
\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)=\int\left[\int_{\underline{I}_{\alpha,h}(x,I;\mathsf{b})}^{\overline{I}_{\alpha,h}(x,I;\mathsf{b})}P\left(x,t\right)P\left(x,t\right)'K\left(t\right)g\left(\Psi\left(t|x,\mathsf{b}\right),x,I\right)dt\right]dx.
$$

The Dominated Convergence Theorem and Proposition C.1-(i)[1], $s\geq 1$, yield that $\overline{\mathsf{R}}^{(2)}(\cdot;\alpha,I)$ is continuously differentiable over $\underline{\mathcal{BI}}_{\alpha,h}$ with, by the Liebniz integral rule,

$$
\begin{aligned}
\overline{\mathsf{R}}^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \overline{\mathsf{R}}_0^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]+\overline{\mathsf{R}}_1^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]-\overline{\mathsf{R}}_2^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}], \\
\overline{\mathsf{R}}_0^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \int_{\mathcal{X}}\left[\int_{\underline{I}_{\alpha,h}(x,I;\mathsf{b})}^{\overline{I}_{\alpha,h}(x,I;\mathsf{b})}P\left(x,t\right)P\left(x,t\right)'K\left(t\right)g^{(1)}\left(\Psi\left(t|x,\mathsf{b}\right),x,I\right)\left[\mathsf{d}'P\left(x,t\right)\right]dt\right]dx, \\
\overline{\mathsf{R}}_1^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \int_{\mathcal{X}}P\left(x,\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)P\left(x,\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)'K\left(\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right) \\
&\qquad\times g\left(\Psi\left(\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)|x,\mathsf{b}\right),x,I\right)\left[\mathsf{d}'\frac{\partial\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)}{\partial\mathsf{b}'}\right]dx, \\
\overline{\mathsf{R}}_2^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}] &= \int_{\mathcal{X}}P\left(x,\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)P\left(x,\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right)'K\left(\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)\right) \\
&\qquad\times g\left(\Psi\left(\overline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)|x,\mathsf{b}\right),x,I\right)\left[\mathsf{d}'\frac{\partial\underline{I}_{\alpha,h}\left(x,I;\mathsf{b}\right)}{\partial\mathsf{b}'}\right]dx.
\end{aligned}
$$

[1] which implies that $g\left(\cdot|\cdot,I\right)$ is bounded away from 0 and infinity.

Proposition C.1-(i) and Assumption R-(i) imply

$$\left\|\overline{\mathsf{R}}_0^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]\right\| \preceq C\max_{x\in\mathcal{X}}\|P(x)\|\,\|\mathsf{d}\| \leq Ch^{-D_{\mathcal{M}}/2}\|\mathsf{d}\|.$$

The operators $\overline{\mathsf{R}}_i^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]$, $i=1,2$, can be studied in a similar way so that only $i=1$ is considered. Observe

$$\frac{\partial\overline{I}_{\alpha,h}(x,I;\mathsf{b})}{\partial\mathsf{b}'} = \begin{cases} 0 & \text{if } \overline{I}_{\alpha,h} \leq \Delta[B(1|x,I)|x,\mathsf{b}] \\ \frac{\partial\Delta[B(1|x,I)|x,\mathsf{b}]}{\partial\mathsf{b}'} = -\frac{P(x,\Delta(B(1|x,I)|x,\mathsf{b}))}{\Psi^{(1)}(\Delta(B(1|x,I)|x,\mathsf{b})|x,\mathsf{b})} & \text{if } \overline{I}_{\alpha,h} > \Delta[B(1|x,I)|x,\mathsf{b}] \end{cases}.$$

But, for $h$ small enough,

$$\begin{aligned}\Delta[B(1|x,I)|x,\mathsf{b}] &= \frac{\Phi[B(1|x,I)|x,\mathsf{b}]-\alpha}{h} = \frac{\min\left\{\alpha+h\overline{I}_{\alpha,h},\Phi[B(1|x,I)|x,\mathsf{b}]\right\}-\alpha}{h}\\ &\geq \frac{\min\left\{\alpha+h\overline{I}_{\alpha,h},\Phi[B(1|x,I)|x,\mathsf{b}^*(\alpha|I)]-Ch^{-D_{\mathcal{M}}/2}\|\mathsf{b}-\mathsf{b}^*(\alpha|I)\|\right\}-\alpha}{h}\\ &\geq \frac{\min\left\{\alpha+h\overline{I}_{\alpha,h},G[B(1|x,I)|x,I]-Ch^{s+1}-Ch\right\}-\alpha}{h}\\ &\geq \frac{\min\left\{\alpha+h\min\left(\frac{1-\alpha}{h},1\right),1-Ch\right\}-\alpha}{h}\end{aligned}$$

uniformly in $\alpha$, $x$ and $\mathsf{b}$ in $\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)$ by Lemma B.1. Hence, if $\alpha\leq 1-C'h$ with $C'\geq 1$ large enough

$$\Delta[B(1|x,I)|x,\mathsf{b}] \geq \frac{\min\{\alpha+h,1-Ch\}-\alpha}{h} \geq 1 \geq \overline{I}_{\alpha,h}$$

so that $\frac{\partial \overline{I}_{\alpha,h}(x,I;\mathsf{b})}{\partial \mathsf{b}'} = 0$. Hence since $\mathcal{B}\left(\mathsf{b}^*(\alpha|I), Ch^{D_{\mathcal{M}}/2+1}\right) \subset \underline{\mathcal{BI}}_{\alpha,h}$ and by definition of $\underline{\mathcal{BI}}_{\alpha,h}$

$$\begin{aligned}
\left\|\overline{\mathsf{R}}_1^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]\right\| &\leq C\mathbb{I}[\alpha \geq 1 - C'h] \\
&\quad \times \left\| \int_{\mathcal{X}} P\left(x, \overline{I}_{\alpha,h}(x,I;\mathsf{b})\right) P\left(x, \overline{I}_{\alpha,h}(x,I;\mathsf{b})\right)' \frac{\mathsf{d}' P(x, \Delta(B(1|x,I)|x,\mathsf{b}))}{\Psi(\Delta(B(1|x,I)|x,\mathsf{b})|x,\mathsf{b})} dx \right\| \\
&\leq Ch^{-1}\mathbb{I}[\alpha \geq 1 - C'h] \max_{x\in\mathcal{X}} \|P(x)\| \|\mathsf{d}\| \leq Ch^{-1}h^{-D_{\mathcal{M}}/2} \|\mathsf{d}\| \mathbb{I}[\alpha \geq 1 - C'h] \\
&\leq C\frac{h^{-D_{\mathcal{M}}/2}}{\alpha(1-\alpha)+h} \|\mathsf{d}\|.
\end{aligned}$$

Substituting in the expression of $\overline{\mathsf{R}}^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]$ then gives uniformly in $\mathsf{d}$

$$\max_{\alpha\in[0,1]} \max_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I), Ch^{D_{\mathcal{M}}/2+1}\right)} (\alpha(1-\alpha)+h) \left\|\overline{\mathsf{R}}^{(3)}(\mathsf{b};\alpha,I)[\mathsf{d}]\right\| \leq Ch^{-D_{\mathcal{M}}/2} \|\mathsf{d}\|.$$

The Taylor inequality shows that (i) holds.

For (ii), the expression of $\overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)$, Assumptions A and R-(i), Proposition C.1-(i), which imply that the eigenvalues of $\int P(x) P'(x) g[B(\alpha|x,I), x, I] dx$ stay bounded away 0 and infinity, Lemma B.1-(iii) and Proposition C.1-(i) give that, uniformly in $\alpha$ and $x$

$$\begin{aligned}
\overline{I}_{\alpha,h}[x,I;\mathsf{b}^*(\alpha|I)] &= \overline{I}_{\alpha,h} \wedge \frac{\Phi[B(1|x,I)|x,\mathsf{b}^*(\alpha|I)] - \alpha}{h} \\
&= \overline{I}_{\alpha,h} \wedge \frac{1 + o(h^{s+1}) - \alpha}{h} = \overline{I}_{\alpha,h} + o(h^s), \\
\underline{I}_{\alpha,h}[x,I;\mathsf{b}^*(\alpha|I)] &= \underline{I}_{\alpha,h} + o(h^s),
\end{aligned}$$

$$
\begin{aligned}
\overline{\mathsf{R}}^{(2)}\left(\mathsf{b}^{*}(\alpha|I);\alpha,I\right) &= \int\left[\int_{\underline{I}_{\alpha,h}[x,I;\mathsf{b}^{*}(\alpha|I)]}^{\overline{I}_{\alpha,h}[x,I;\mathsf{b}^{*}(\alpha|I)]}\pi(t)\,\pi(t)'\,K(t)\,g\left(\Psi\left(t|x,\mathsf{b}^{*}(\alpha|I)\right)|x,I\right)dt\right] \\
&\qquad \otimes P(x)\,P(x)'\,f(x,I)\,dx \\
&= \int\left[\int_{\underline{I}_{\alpha,h}+o(h^{s})}^{\overline{I}_{\alpha,h}+o(h^{s})}\pi(t)\,\pi(t)'\,K(t)\,g\left[B(\alpha+ht|x,I)+o\left(h^{s+1}\right)|x,I\right]dt\right] \\
&\qquad \otimes P(x)\,P(x)'\,f(x,I)\,dx \\
&= \int\left[\int_{\underline{I}_{\alpha,h}+o(h^{s})}^{\overline{I}_{\alpha,h}+o(h^{s})}\pi(t)\,\pi(t)'\,K(t)\left(\frac{1}{B^{(1)}(\alpha+ht|x,I)}+o\left(h^{s+1}\right)\right)dt\right] \\
&\qquad \otimes P(x)\,P(x)'\,f(x,I)\,dx \\
&= \int\left[\int_{\underline{I}_{\alpha,h}+o(h^{s})}^{\overline{I}_{\alpha,h}+o(h^{s})}\pi(t)\,\pi(t)'\,K(t)\left(\frac{1}{B^{(1)}(\alpha|x,I)}-ht\frac{B^{(2)}(\alpha|x,I)}{\left(B^{(1)}(\alpha|x,I)\right)^{2}}+o(h)\right)dt\right] \\
&\qquad \otimes P(x)\,P(x)'\,f(x,I)\,dx \\
&= \int\Omega_{h}(\alpha)\otimes\frac{P(x)\,P(x)'}{B^{(1)}(\alpha|x,I)}f(x,I)\,dx \\
&\qquad -h\int\Omega_{1h}(\alpha)\otimes\frac{P(x)\,P(x)'\,B^{(2)}(\alpha|x,I)}{\left(B^{(1)}(\alpha|x,I)\right)^{2}}f(x,I)\,dx+o(h)
\end{aligned}
$$

where the last $o(h)$ term is with respect of the matrix norm. This together the fact that the eigenvalues of the matrices $\Omega_h(\alpha)$ and $\int_{\mathcal{X}} P(x)\,P(x)'\,dx$ are bounded away from 0 and infinity, the fact that $B^{(1)}(\alpha|x,I)$ is bounded away from 0 and infinity shows that (ii) holds.$\square$

**Proof of Lemma C.3.** Write $A_{\alpha,h}^{-1} = D_{\alpha,h} + B_{\alpha,h}$ where $D_{\alpha,h}$ is the diagonal of $A_{\alpha,h}^{-1}$ and $B_{\alpha,h} = A_{\alpha,h}^{-1} - D_{\alpha,h}$. Provided the series converges

$$
A_{\alpha,h} = D_{\alpha,h}^{-1/2}\left\{\sum_{n=0}^{\infty}\left(D_{\alpha,h}^{-1/2}B_{\alpha,h}D_{\alpha,h}^{-1/2}\right)^{n}\right\}D_{\alpha,h}^{-1/2}.
$$

Proposition C.1-(i) and Assumption R-(i) ensure that the entries of $D_{\alpha,h}^{-1/2}$ are bounded in absolute value by $C < \infty$ for all $\alpha$ and $L$. It also gives

$$\left| \frac{\mathbb{E}\left[\frac{\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|x_\ell,I_\ell)}\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P_{k_1}(x_\ell)\,\pi_{p_1}(t)\,P_{k_2}(x_\ell)\,\pi_{p_2}(t)\,K(t)\,dt\right]}{\mathbb{E}^{1/2}\left[\frac{\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|x_\ell,I_\ell)}\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P_{k_1}^2(x_\ell)\,\pi_{p_1}^2(t)\,K(t)\,dt\right]\mathbb{E}^{1/2}\left[\frac{\mathbb{I}(I_\ell=I)}{B^{(1)}(\alpha|x_\ell,I_\ell)}\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} P_{k_2}^2(x_\ell)\,\pi_{p_2}^2(t)\,K(t)\,dt\right]} \right| \leq \varrho < 1$$

for all $1 \leq k_1, k_2 \leq K$ and $0 \leq p_1, p_2 \leq s+1$, that is all the entries of $D_{\alpha,h}^{-1/2}B_{\alpha,h}D_{\alpha,h}^{-1/2}$ are bounded by $\varrho$ in absolute value. By Assumption R-(ii), the entries of $D_{\alpha,h}^{-1/2}B_{\alpha,h}D_{\alpha,h}^{-1/2}$ are bounded by the ones of $\varrho\,\mathrm{Id}\otimes(T'+T)$, where $T$ is a lower $c/2$ band matrix with band entries equal to 1 and Id is the $(s+2)\times(s+2)$ identity matrix. Hence the absolute value of the entries of $A_{\alpha,h}$ are bounded by the entries of

$$C\,\mathrm{Id}\otimes\left(\sum_{n=\infty}^{\infty}\varrho^n\left(T^{n'}+T^n\right)\right).$$

Since $T$ is a triangular $c-$band nilpotent matrix, it follows that $|A_{\alpha,h}(j_1,j_2)| \leq C\rho^{|j_2-j_1|}$ with $0 < \varrho \leq \rho < 1$, for all $\alpha$ and $L$. It follows

$$\max_L \max_{\alpha\in[0,1]} \max_{1\leq j_1\leq(s+1)K} \sum_{j_2=1}^{(s+1)K} |A_{\alpha,h}(j_1,j_2)| \leq C\sum_n \rho^n < \infty$$

which ends the proof of the Lemma. □

## F.2 Lemmas B.3, B.4 and B.5

The proofs of the lemmas grouped here make use of a deviation inequality from Massart (2007). Consider $n$ independent random variables $Z_\ell$ and, for a known real function $\xi(z,\theta)$ separable with respect to $\theta\in\Theta$, $Z_\ell(\theta)=\xi(Z_\ell,\theta)$ where $\theta$ is a parameter. Let $\underline{\xi}(\cdot)\leq\overline{\xi}(\cdot)$ be two functions. A *bracket* $\left[\underline{\xi},\overline{\xi}\right]$ is the set of all functions $\xi(\cdot)$ such that $\underline{\xi}(z)\leq\xi(z)\leq\overline{\xi}(z)$ for all $z$. The next proposition follows from Massart (2007, Theorem 6.8 and Corollary 6.9).

**Proposition F.1** *Assume that* $\sup_{\theta\in\Theta}|Z_\ell(\theta)| \leq M_\infty$*,* $\sup_{\theta\in\Theta}\mathrm{Var}(Z_\ell(\theta)) \leq M_2^2$ *for all* $\ell$ *and that for any* $\epsilon > 0$ *there exists brackets* $\left[\underline{\xi}_j, \overline{\xi}_j\right] \subset [-b, b]$*,* $j = 1, \ldots, \exp(H(\epsilon))$*, such that*

$$\mathbb{E}\left[\left(\overline{\xi}_j(Z_i) - \underline{\xi}_j(Z_i)\right)^2\right] \leq \frac{\epsilon^2}{2} \text{ and } \{\xi(z,\theta), \theta\in\Theta\} \subset \bigcup_{j=1}^{\exp(H(\epsilon))}\left[\underline{\xi}_j, \overline{\xi}_j\right].$$

*Let*

$$\mathcal{H}_L = 54\int_0^{M_2/2}\sqrt{\min(L, H(\epsilon))}d\epsilon + \frac{2(M_\infty + M_2)H(M_2)}{L^{1/2}}.$$

*Then, for any* $t \in \left[0, 10L^{1/2}M_2/M_\infty\right]$*,*

$$\mathbb{P}\left(\sup_{\theta\in\Theta}\left|\sum_{i=1}^n\{Z_\ell(\theta) - \mathbb{E}[Z_\ell(\theta)]\}\right| \geq L^{1/2}\{\mathcal{H}_L + t\}\right) \leq 2\exp\left(-\frac{t^2}{25}\right).$$

**Proof of Lemma B.3.** Note that $\widehat{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)$ is a $c(s+2)$-band matrix, so that the order of its matrix norm is the same than the order of its largest entry. The generic entry of $\widehat{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I) - \overline{\mathsf{R}}^{(2)}(\mathsf{b};\alpha,I)$ can be written as

$$\widehat{\mathsf{r}}(\mathsf{b};\alpha,I) = \frac{1}{Lh^{(D_{\mathcal{M}}+1)/2}}\sum_{\ell=1}^L\xi_\ell(\mathsf{b};\alpha)$$

where the $\xi_\ell(\mathsf{b};\alpha)$ are centered iid with

$$\begin{aligned}
\xi_\ell(\mathsf{b};\alpha) = \sum_{i=1}^{I_\ell}&\{\mathbb{I}[B_{i\ell}\in\Psi(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}), I_\ell = I]\xi_{i\ell}(\mathsf{b}) \\
&-\mathbb{E}[\mathbb{I}[B_{i\ell}\in\Psi(\mathcal{I}_{\alpha,h}|x_\ell,\mathsf{b}), I_\ell = I]\xi_{i\ell}(\mathsf{b})]\} \\
\xi_{i\ell}(\mathsf{b}) = \frac{h^{D_{\mathcal{M}}/2}}{h^{1/2}}&\frac{P_{k_1}(x_\ell)P_{k_2}(x_\ell)}{\Psi^{(1)}(\Delta(B_{i\ell}|x_\ell,\mathsf{b})|x_\ell,\mathsf{b})/h}K_p(\Delta(B_{i\ell}|x_\ell,\mathsf{b})), \\
K_p(\Delta(B_{i\ell}|x_\ell,\mathsf{b})) = &\frac{\Delta^{p_1+p_2}(B_{i\ell}|x_\ell,\mathsf{b})}{p_1!p_2!}K(\Delta(B_{i\ell}|x_\ell,\mathsf{b})).
\end{aligned}$$

The proof of the Lemma follows from Proposition F.1. Observe

$$|\xi_\ell(\mathsf{b};\alpha)| \leq C\frac{h^{D_{\mathcal{M}}/2}\max_{x\in\mathcal{X}}\|P(x)\|^2}{h^{1/2}} \leq M_\infty \text{ with } M_\infty \asymp h^{-(D_{\mathcal{M}}+1)/2}.$$

for all $\alpha$ in $[0,1]$ and all admissible $\mathsf{b}$. For the variance, Lemma B.1-(iii,iv) gives

$$\begin{aligned}
|\Delta(B_{i\ell}|x_\ell,\mathsf{b})| &= \left|\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b})-\alpha}{h}\right| \\
&\leq \left|\frac{G(B_{i\ell}|x_\ell,I_\ell)-\alpha}{h}\right| + \left|\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b}^*(\alpha|I_\ell))-G(B_{i\ell}|x_\ell,\mathsf{b})}{h}\right| \\
&+ \left|\frac{\Phi(B_{i\ell}|x_\ell,\mathsf{b})-\Phi(B_{i\ell}|x_\ell,\mathsf{b}^*(\alpha|I_\ell))}{h}\right| \\
&\leq \left|\frac{G(B_{i\ell}|x_\ell,I_\ell)-\alpha}{h}\right| + o(h^s) + O\left(\frac{h^{-D_{\mathcal{M}}/2}\times h^{D_{\mathcal{M}}/2+1}}{h}\right) \\
&= \left|\frac{G(B_{i\ell}|x_\ell,I_\ell)-\alpha}{h}\right| + O(1)
\end{aligned}$$

uniformly. It follows that, $U_{i\ell} = G(B_{i\ell}|x_\ell,I_\ell)$ being a uniform random variable independent of $(x_\ell,I_\ell)$

$$\begin{aligned}
\operatorname{Var}(\xi_\ell(\mathsf{b};\alpha)) &\leq CI^2h^{D_{\mathcal{M}}}\max_{x\in\mathcal{X}}\|P(x)\|^2\int_{\mathcal{X}}|P_{k_1}(x)P_{k_2}(x)|\,dx\int\mathbb{I}_{[-C,C]}\left(\frac{u-\alpha}{h}\right)\frac{du}{h} \\
&\leq CI^2h^{D_{\mathcal{M}}}\max_{x\in\mathcal{X}}\|P(x)\|^2\left(\int_{\mathcal{X}}P_{k_1}^2(x)\,dx\right)^{1/2}\left(\int_{\mathcal{X}}P_{k_2}^2(x)\,dx\right)^{1/2} \\
&\leq M_2^2 \text{ with } M_2<\infty
\end{aligned}$$

under Assumption R, uniformly in $\mathsf{b}$ and $\alpha$.

Consider now the brackets covering. The key observation is that $\xi_\ell(\mathsf{b};\alpha)$ only depends on a finite dimension subvector of $\mathsf{b}$, $\mathsf{b}^{(k_1,k_2)}$ which groups the entries of $\mathsf{b}$ corresponding to those $P_k(\cdot)$ such that $P_k(\cdot)P_{k_1}(\cdot)\neq 0$ or $P_k(\cdot)P_{k_2}(\cdot)\neq 0$, so that the dimension of $\mathsf{b}^{(k_1,k_2)}$ is less than $c(s+2)$ under Assumption R-(ii). Consequently the class to be bracketed is

$$\mathcal{F}=\left\{\xi_\ell\left(\mathsf{b}^{(k_1,k_2)};\alpha\right);\alpha\in[0,1],\mathsf{b}^{(k_1,k_2)}\in\mathcal{B}\left(\mathsf{b}^{(k_1,k_2)*}(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)\right\}.$$

Lemma B.1-(iii), $1/\left(Lh^{D_{\mathcal{M}}+1}\right)=o(1)$, van de Geer (1999, p.20) and arguing as Guerre and Sabbah (2012, 2014) imply that $\mathcal{F}$ can be bracketed with a number of brackets

$$\exp\left(H_L(\epsilon)\right) \asymp \left(\frac{L^C}{\epsilon}\right)^C$$

so that

$$\int_0^{M_2/2} \sqrt{\min\left(L, H_L(\epsilon)\right)}d\epsilon \leq \left(\frac{M_2}{2}\right)^{1/2}\left(\int_0^{M_2/2} H_L(\epsilon)\, d\epsilon\right)^{1/2} = O\left(\log L\right)^{1/2}$$

and for the item $\mathcal{H}_L$ of Proposition F.1,

$$\mathcal{H}_L = O\left(\log L\right)^{1/2} + O\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2} = O\left(\log L\right)^{1/2}$$

since $1/\left(Lh^{D_{\mathcal{M}}+1}\right)$ is bounded. Hence, by Proposition F.1 for $t \leq 10L^{1/2}M_2/M_\infty$ diverges

$$\begin{aligned}\mathbb{P}&\left(\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}\sup_{\alpha\in[0,1]}\sup_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)}\left|\widehat{\mathsf{r}}\left(\mathsf{b};\alpha,I\right)\right| \geq C\log^{1/2}L+t\right)\\ &\leq 2\exp\left(-\frac{t^2}{25}\right)\end{aligned}$$

uniformly over all the non zero entries $\widehat{\mathsf{r}}\left(\mathsf{b};\alpha,I\right)$ of the band matrix $\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)-\overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)$. This gives, by the Bonferroni inequality

$$\begin{aligned}\mathbb{P}&\left(\sup_{\alpha\in[0,1]}\sup_{\mathsf{b}\in\mathcal{B}\left(\mathsf{b}^*(\alpha|I),Ch^{D_{\mathcal{M}}/2+1}\right)}\left\|\widehat{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)-\overline{\mathsf{R}}^{(2)}\left(\mathsf{b};\alpha,I\right)\right\| \geq \frac{C\log^{1/2}L+t}{\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}}\right)\\ &\leq CK\exp\left(-\frac{t^2}{25}\right)\end{aligned}$$

which implies the result of the lemma since $t \leq 10L^{1/2}M_2/M_\infty = O\left(Lh^{D_{\mathcal{M}}+1}\right)^{1/2}$ can be set to $t=\tau\log^{1/2}L$ for an arbitrary large $\tau$ as $\log L/\left(Lh^{D_{\mathcal{M}}+1}\right)=o(1)$. $\square$

**Proof of Lemma B.4.** The proof of Lemma B.4 is similar to the one of Lemma B.3. The generic entry of $\widehat{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)-\overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)$ writes

$$\widehat{\mathsf{r}}(\mathsf{b};\alpha,I)=\frac{1}{L}\sum_{\ell=1}^{L}\xi_\ell(\mathsf{b};\alpha)$$

where the $\xi_\ell(\mathsf{b};\alpha)$ are centered iid with, for $K_p(t)=t^pK(t)/p!$,

$$\xi_\ell(\mathsf{b};\alpha)=\sum_{i=1}^{I_\ell}\left(\mathbb{I}(I_\ell=I)\,\xi_{i\ell}(\mathsf{b};\alpha)-\mathbb{E}\left[\mathbb{I}(I_\ell=I)\,\xi_{i\ell}(\mathsf{b};\alpha)\right]\right),$$
$$\xi_{i\ell}(\mathsf{b};\alpha)=P_k(x_\ell)\left\{\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\mathbb{I}\left[B_{i\ell}\leq\Psi(t|x_\ell,\mathsf{b})\right]-(\alpha+ht)\right\}K_p(t)\,dt\right\}.$$

This gives

$$\left|\frac{\xi_\ell(\mathsf{b};\alpha)}{(h+\alpha(1-\alpha))^{1/2}}\right|\leq Ch^{-1/2}\max_{x\in\mathcal{X}}\|P(x)\|\leq M_\infty \text{ with } M_\infty\asymp h^{-(D_{\mathcal{M}}+1)/2}.$$

For the computation of the variance, Lemma B.1-(iii,iv) and Proposition C.1-(i) give uniformly in $\alpha$, $t$ in $\mathcal{I}_{\alpha,h}$ the admissible $\mathsf{b}$ and $x_\ell$, and for the uniform $U_{i\ell}=G(B_{i\ell}|x_\ell,I_\ell)$,

$$\begin{aligned}\mathbb{I}\left[B_{i\ell}\leq\Psi(t|x_\ell,\mathsf{b})\right]&=\mathbb{I}\left[B_{i\ell}\leq\Psi(t|x_\ell,\mathsf{b}^*(\alpha|I))+O(h)\right]\\&=\mathbb{I}\left[B(U_{i\ell}|x_\ell,I_\ell)\leq B(\alpha+ht|x_\ell,I_\ell)+O(h)\right]\\&=\mathbb{I}\left[U_{i\ell}\leq G\left[B(\alpha+ht|x_\ell,I_\ell)+O(h)\,|x_\ell,I_\ell\right]\right]\\&=\mathbb{I}\left[U_{i\ell}\leq\alpha+ht+O(h)\right].\end{aligned}$$

It then follows, since $U_{i\ell}$ is independent of $(x_\ell, I_\ell)$

$$\begin{aligned}
&\mathbb{E}\left[\xi_{i\ell}^2(\mathsf{b};\alpha)\,|I_\ell\right] \\
&\quad\leq \mathbb{E}\left[P_k^2(x_\ell)\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\mathbb{I}\left[U_{i\ell}\leq\alpha+h\,(t_1\wedge t_2)+O\,(h)\right]K_p(t_1)\,K_p(t_2)\,dt_1dt_2|I_\ell\right] \\
&\qquad-2\mathbb{E}\left[P_k^2(x_\ell)\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\mathbb{I}\left[U_{i\ell}\leq\alpha+ht_1+O\,(h)\right](\alpha+ht_2)\,K_p(t_1)\,K_p(t_2)\,dt_1dt_2|I_\ell\right] \\
&\qquad+\mathbb{E}\left[P_k^2(x_\ell)\,|I_\ell\right]\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}(\alpha+ht_1)\,(\alpha+ht_2)\,K_p(t_1)\,K_p(t_2)\,dt_1dt_2 \\
&\quad=\mathbb{E}\left[P_k^2(x_\ell)\,|I_\ell\right]\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\int_{\underline{\mathsf{I}}_{\alpha,\mathsf{h}}}^{\bar{\mathsf{I}}_{\alpha,\mathsf{h}}}\left\{\alpha+O\,(h)-\alpha^2\right\}K_p(t_1)\,K_p(t_2)\,dt_1dt_2\leq C\,(h+\alpha\,(1-\alpha))
\end{aligned}$$

uniformly in $\alpha$ and $\mathsf{b}$. Hence, uniformly in $\alpha$ and $\mathsf{b}$

$$\operatorname{Var}\left(\frac{\xi_\ell(\mathsf{b};\alpha)}{(h+\alpha\,(1-\alpha))^{1/2}}\right)\leq M_2^2 \text{ with } M_2<\infty.$$

The bracketing part of the proof is similar to the one of Lemma B.3 and gives

$$\mathcal{H}_L=O\,(\log L)^{1/2}+O\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}=O\,(\log L)^{1/2}.$$

Arguing with Proposition F.1 then shows that the order of the largest entry in $\widehat{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)-\overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)$ is $O_{\mathbb{P}}(\log L/L)^{1/2}$, which gives uniformly

$$\left\|\widehat{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)-\overline{\mathsf{R}}^{(1)}(\mathsf{b};\alpha,I)\right\|=K^{1/2}O_{\mathbb{P}}\left(\frac{\log L}{L}\right)^{1/2}=O_{\mathbb{P}}\left(\frac{\log L}{Lh^{D_{\mathcal{M}}}}\right)^{1/2}$$

and the Lemma is proved. □

**Proof of Lemma B.5.** For (i), define

$$\mathbf{P} = \mathbb{E}\left[\mathbb{I}(I_\ell = I)\, P(x_\ell)\, P(x_\ell)'\right],$$
$$\mathbf{P}_0 = \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)\, P(x_\ell)\, P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right],$$
$$\mathbf{P}_1 = \mathbb{E}\left[\frac{\mathbb{I}(I_\ell = I)\, B^{(2)}(\alpha|x_\ell, I_\ell)\, P(x_\ell)\, P(x_\ell)'}{\left(B^{(1)}(\alpha|x_\ell, I_\ell)\right)^2}\right],$$

and abbreviate $\Omega_h(\alpha)$, $\Omega_{1h}(\alpha)$ in $\Omega$, $\Omega_1$. It holds

$$\operatorname{Var}\left(\widehat{\mathsf{e}}(\alpha|I)\right) = \left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right]^{-1} \operatorname{Var}\left[\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right]\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right]^{-1}$$

with by Lemma B.2

$$\begin{aligned}
\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right]^{-1} &= \left[\Omega \otimes \mathbf{P}_0 - h\Omega_1 \otimes \mathbf{P}_1 + o(h)\right]^{-1} \\
&= \left[\mathrm{Id} - h\left(\Omega^{-1}\Omega_1\right) \otimes \left(\mathbf{P}_0^{-1}\mathbf{P}_1\right) + o(h)\right]^{-1} \Omega^{-1} \otimes \mathbf{P}_0^{-1} \\
&= \Omega^{-1} \otimes \mathbf{P}_0^{-1} + h\left(\Omega^{-1}\Omega_1\Omega^{-1}\right) \otimes \left(\mathbf{P}_0^{-1}\mathbf{P}_1\mathbf{P}_0^{-1}\right) + o(h)
\end{aligned}$$

uniformly in $\alpha$ where the remainder term $o(h)$ is with respect to the matrix norm. For $\operatorname{Var}\left[\widehat{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha, I\right)\right]$, define

$$\omega_0 = \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \pi(t)\, K(t)\, dt, \quad \omega_1 = \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \pi(t)\, K(t)\, dt,$$
$$\mathbf{\Pi}_m = \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}} \min(t_1, t_2)\, \pi(t_1)\, \pi(t_2)'\, K(t_1)\, K(t_2)\, dt.$$

Now (C.3) in the proof of Theorem C.4 and Lemma B.1-(iii,iv) show that $(LI)\,\mathrm{Var}\left[\widehat{\mathsf{R}}^{(1)}\left(\bar{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]$ admits the expansion, with uniform remainder terms,

$$\begin{aligned}
&\mathbb{E}\left[\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\{G\left[B\left(\alpha+ht_1|x_\ell,I_\ell\right)\wedge B\left(\alpha+ht_2|x_\ell,I_\ell\right)+o\left(h\right)|x_\ell,I_\ell\right]\right.\\
&\qquad -G\left[B\left(\alpha+ht_1|x_\ell,I_\ell\right)+o\left(h\right)|x_\ell,I_\ell\right]\left(\alpha+ht_2\right)-G\left[B\left(\alpha+ht_2|x_\ell,I_\ell\right)+o\left(h\right)|x_\ell,I_\ell\right]\left(\alpha+ht_1\right)\\
&\qquad \left.+\left(\alpha+ht_1\right)\left(\alpha+ht_2\right)\}\,\pi\left(t_1\right)\pi\left(t_2\right)'K\left(t_1\right)K\left(t_2\right)dt_1dt_2\otimes\mathbb{I}\left(I_\ell=I\right)P\left(x_\ell\right)P\left(x_\ell\right)'\right]\\
&\quad=\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\int_{\underline{I}_{\alpha,h}}^{\bar{I}_{\alpha,h}}\left\{\alpha+h\left(t_1\wedge t_2\right)-\alpha^2-h\alpha\left(t_1+t_2\right)\right\}\pi\left(t_1\right)\pi\left(t_2\right)'K\left(t_1\right)K\left(t_2\right)dt_1dt_2+o\left(h\right)\\
&\quad=\alpha\left(1-\alpha\right)\omega_0\omega_0'\otimes\mathbf{P}+h\left\{\mathbf{\Pi}_m-\alpha\left(\omega_0\omega_1'+\omega_1\omega_0'\right)\right\}\otimes\mathbf{P}+o\left(h\right).
\end{aligned}$$

Hence an elementary expansion gives, uniformly in $\alpha\in[0,1]$, $\mathrm{Var}\left(\widehat{\mathsf{e}}\left(\alpha|I\right)\right)=\mathsf{V}_e/\left(LI\right)+o\left(h\right)$ with

$$\begin{aligned}
\mathsf{V}_e&=\alpha\left(1-\alpha\right)\left[\Omega^{-1}\omega_0\omega_0'\Omega^{-1}\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right]\\
&\quad+h\alpha\left(1-\alpha\right)\left[\Omega^{-1}\Omega_1\Omega^{-1}\omega_0\omega_0'\Omega^{-1}\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}_1\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right]\\
&\quad+h\alpha\left(1-\alpha\right)\left[\Omega^{-1}\omega_0\omega_0'\Omega^{-1}\Omega_1\Omega^{-1}\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\mathbf{P}_1\mathbf{P}_0^{-1}\right]\\
&\quad+h\left[\Omega^{-1}\left(\mathbf{\Pi}_m-\left(\omega_1\omega_0'+\omega_0\omega_1'\right)\right)\Omega^{-1}\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right].
\end{aligned}$$

Observe now that $\Omega^{-1}\omega_0=s_0$, $\Omega^{-1}\omega_1=s_1$ and $\Omega^{-1}\Omega_1\Omega^{-1}\omega_0=\Omega^{-1}\Omega_1 s_0=\Omega^{-1}\omega_1=s_1$. This gives

$$\begin{aligned}
\mathsf{V}_e&=\alpha\left(1-\alpha\right)\left[s_0's_0\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right]\\
&\quad+h\alpha\left(1-\alpha\right)\left[s_1's_0\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}_1\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right]\\
&\quad+h\alpha\left(1-\alpha\right)\left[s_0's_1\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\mathbf{P}_1\mathbf{P}_0^{-1}\right]\\
&\quad+h\left[\Omega^{-1}\mathbf{\Pi}_m\Omega^{-1}-\left(s_1s_0'+s_0s_1'\right)\right]\otimes\left[\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right].
\end{aligned}$$

Since the eigenvalues of $\mathbf{P}_0^{-1}$, $\mathbf{P}$, $\mathbf{P}_1$, $\Omega^{-1}$ and $\Omega_1$ are bounded away from infinity uniformly in $\alpha$, it follows that $\max_{\alpha\in[0,1]} \|\mathrm{Var}(\widehat{\mathsf{e}}_0(\alpha|I))\| = O(1/L)$ and then

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}} \mathrm{Var}\left(P(x)'\widehat{\mathsf{e}}_0(\alpha|I)\right) = O\left(\frac{\max_{x\in\mathcal{X}}\|P(x)\|^2}{L}\right) = O\left(\frac{1}{Lh^{D_{\mathcal{M}}}}\right).$$

For $\mathrm{Var}(\widehat{\mathsf{e}}_1(\alpha|I)/h)$, observe that $\widehat{\mathsf{e}}_1(\alpha|I) = S_1\widehat{\mathsf{e}}(\alpha|I)$ with

$$S_1 = s_1' \otimes \mathrm{Id}$$

it holds

$$\begin{aligned} S_1\mathsf{V}_e S_1' &= h\left(s_1'\Omega^{-1}\mathbf{\Pi}_m\Omega^{-1}s_1\right)\left(\mathbf{P}_0^{-1}\mathbf{P}\mathbf{P}_0^{-1}\right) \\ &= hv_h^2(\alpha)\,\mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)\,P(x_\ell)\,P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] \\ &\qquad \times \mathbb{E}\left[\mathbb{I}(I_\ell = I)\,P(x_\ell)\,P(x_\ell)'\right]\mathbb{E}^{-1}\left[\frac{\mathbb{I}(I_\ell = I)\,P(x_\ell)\,P(x_\ell)'}{B^{(1)}(\alpha|x_\ell, I_\ell)}\right] \end{aligned}$$

as $v_h^2(\alpha) = s_1'\Omega^{-1}\mathbf{\Pi}_m\Omega^{-1}s_1$. This gives the result for $\mathrm{Var}(\widehat{\mathsf{e}}_1(\alpha|I)/h)$ and $\mathrm{Var}\left(P(x)'\widehat{\mathsf{e}}_1(\alpha|I)/h\right)$.

For (ii), we just show that $\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|P(x)'\widehat{\mathsf{e}}_1(\alpha|I)/h\right| = O_{\mathbb{P}}\left(\left(\log L/Lh^{D_{\mathcal{M}}+1}\right)^{1/2}\right)$. Since $\max_{x\in[0,1]}\|P(x)\| = O\left(h^{-D_{\mathcal{M}}/2}\right)$ and

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{P(x)'\widehat{\mathsf{e}}_1(\alpha|I)}{h}\right| \leq \left(\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{P(x)'\widehat{\mathsf{e}}_1(\alpha|I)}{h^{1/2}(1+\|P(x)\|)}\right|\right)\times h^{-1/2}\left(1+\max_{x\in[0,1]}\|P(x)\|\right)$$

it is sufficient to show

$$\max_{(\alpha,x)\in[0,1]\times\mathcal{X}}\left|\frac{P(x)'\widehat{\mathsf{e}}_1(\alpha|I)}{h^{1/2}(1+\|P(x)\|)}\right| = O_{\mathbb{P}}\left(\left(\frac{\log L}{L}\right)^{1/2}\right). \tag{F.3}$$

Write

$$\frac{P(x)'\widehat{\mathsf{e}}_1(\alpha|I)}{h^{1/2}(1+\|P(x)\|)} = \frac{1}{L}\sum_{\ell=1}^{L}\xi_\ell(\alpha,x)$$

with

$$\xi_\ell(\alpha,x)=\sum_{i=1}^{I_\ell}\left(\mathbb{I}\left(I_\ell=I\right)\xi_{i\ell}(\alpha,x)-\mathbb{E}\left[\mathbb{I}\left(I_\ell=I\right)\xi_{i\ell}(\alpha,x)\right]\right),$$

$$\xi_{i\ell}(\alpha,x)=\frac{P(x)'S_1\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1}P(x_\ell)}{h^{1/2}\left(1+\|P(x)\|\right)}$$
$$\times\left\{\int_{\underline{I}_{\alpha,h}}^{\overline{I}_{\alpha,h}}\left\{\mathbb{I}\left[B_{i\ell}\leq\Psi\left(t|x_\ell,\overline{\mathsf{b}}(\alpha|I)\right)\right]-(\alpha+ht)\right\}K(t)\,dt\right\}.$$

This gives, for all $(\alpha,x)\in[0,1]$

$$|\xi_\ell(\alpha,x)|\leq Ch^{-1/2}\frac{\left(\max_{x\in\mathcal{X}}\|P(x)\|\right)^2}{1+\max_{x\in\mathcal{X}}\|P(x)\|}\leq M_\infty \text{ with } M_\infty\asymp h^{-(D_{\mathcal{M}}+1)/2},$$
$$\mathrm{Var}\left(\xi_\ell(\alpha,x)\right)\leq C\frac{\left(\max_{x\in\mathcal{X}}\|P(x)\|\right)^2}{\left(1+\max_{x\in\mathcal{X}}\|P(x)\|\right)^2}\leq M_2 \text{ with } M_2\asymp 1.$$

The Implicit Function Theorem and the FOC $\overline{\mathsf{R}}^{(1)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)=0$, Lemma B.2 with (C.3) and $s\geq D_{\mathcal{M}}/2$ give that $\alpha\mapsto\overline{\mathsf{b}}(\alpha|I)$ is $\|\cdot\|$-Lipshitz with a Lipshitz constant of order $L^C$, as $\alpha\mapsto\left[\overline{\mathsf{R}}^{(2)}\left(\overline{\mathsf{b}}(\alpha|I);\alpha,I\right)\right]^{-1}$ and $x\mapsto P(x)/\left(1+\|P(x)\|\right)$. Lemma B.1-(iii), $1/\left(Lh^{D_{\mathcal{M}}+1}\right)=O(1)$, van de Geer (1999, p.20) and arguing as Guerre and Sabbah (2012, 2014) imply that $\{\xi_\ell(\alpha,x);(\alpha,x)\in[0,1]\times\mathcal{X}\}$ can be bracketed with a number of brackets

$$\exp\left(H_L(\epsilon)\right)\asymp\left(\frac{L^C}{\epsilon}\right)^C.$$

Arguing as in the proof of Lemma B.3 gives, for the item $\mathcal{H}_L$ of Proposition F.1,

$$\mathcal{H}_L=O\left(\log L\right)^{1/2}+O\left(\frac{\log L}{Lh^{D_{\mathcal{M}}+1}}\right)^{1/2}=O\left(\log L\right)^{1/2}$$

and then (F.3) holds. □

## F.3 Lemma E.1

The proof of Lemma E.1 is based on the following lemma.

**Lemma F.2** *Let $k_1(\cdot)$ and $k_2(\cdot)$ be two functions over $[0,1]$ with primitives $K_1(\cdot)$ and $K_2(\cdot)$. Then, if $A$ is a random variable with a uniform distribution over $[0,1]$ and for any choice of the primitives $K_1(\cdot)$ and $K_2(\cdot)$,*

$$\begin{aligned}
&\int_0^1 \int_0^1 k_1(a_1) k_2(a_2) [a_1 \wedge a_2 - a_1 a_2] \, da_1 da_2 \\
&\quad = - \int_0^1 k_2(a_2) \left\{ \int_0^{a_2} (K_1(a_1) - \mathbb{E}[K_1(A)]) \, da_1 \right\} da_2
\end{aligned}$$

**Proof of Lemma F.2.** Observe that

$$\begin{aligned}
&\int_0^1 \int_0^1 k_1(a_1) k_2(a_2) [a_1 \wedge a_2 - a_1 a_2] \, da_1 da_2 \\
&= \mathbb{E}\left[ \int_0^1 k_1(a_1) \mathbb{I}[A \leq a_1] \, da_1 \int_0^1 k_2(a_2) \mathbb{I}[A \leq a_2] \, da_1 \right] \\
&\quad - \mathbb{E}\left[ \int_0^1 k_1(a_1) \mathbb{I}[A \leq a_1] \, da_1 \right] \mathbb{E}\left[ \int_0^1 k_2(a_2) \mathbb{I}[A \leq a_2] \, da_1 \right] \\
&= \operatorname{Cov}\left( \int_0^A k_1(a) \, da, \int_0^A k_2(a) \, da \right) = \operatorname{Cov}(K_1(A), K_2(A))
\end{aligned}$$

which does not depend upon the choice of the primitives. Integrating by parts now gives

$$\begin{aligned}
\operatorname{Cov}(K_1(A), K_2(A)) &= \int_0^1 K_2(a_2) (K_1(a_2) - \mathbb{E}[K_2(A)]) \, da_2 \\
&= \int_0^1 K_2(a_2) \, d\left[ \int_0^{a_2} (K_1(a_1) - \mathbb{E}[K_2(A)]) \, da_1 \right] \\
&= - \int_0^1 k_2(a_2) \left\{ \int_0^{a_2} (K_1(a_1) - \mathbb{E}[K_2(A)]) \, da_1 \right\} da_2
\end{aligned}$$

since $\int_0^{a_2} (K_1(a_1) - \mathbb{E}[K_2(A)]) \, da_1$ vanishes for $a_2 = 0$ and $a_2 = 1$. □

**Proof of Lemma E.1** It is assumed that $h < 1/2$ all over the proof. Set $k_h(a_1;\alpha_1) = \frac{1}{h}\pi\left(\frac{a_1-\alpha_1}{h}\right)K\left(\frac{a_1-\alpha_1}{h}\right)$ and $K_h(a_1;\alpha_1) = \int_{-\infty}^{a_1} k_h(a;\alpha_1)\,da$. It follows from Lemma F.2 that

$$
\begin{aligned}
\mathcal{C}_h &= \int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\left\{\int_0^1\int_0^1 k_h(a_2;\alpha_2)\,k_h(a_1;\alpha_1)'\,[a_1\wedge a_2 - a_1a_2]\,da_1da_2\right\}d\alpha_1d\alpha_2\\
&= -\int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\int_0^1 k_h(a_2;\alpha_2)\left\{\int_0^{a_2}\left(K_h(a_1;\alpha_1)-\mathbb{E}\left[K_h(A;\alpha_1)\right]\right)'da_1\right\}da_2\\
&= -\mathcal{I}_h + \mathcal{J}_h
\end{aligned}
$$

with

$$
\begin{aligned}
\mathcal{I}_h &= \int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\int_0^1 k_h(a_2;\alpha_2)\left\{\int_0^{a_2} K_h(a_1;\alpha_1)'\,da_1\right\}da_2\\
&= \int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\int_0^1 \frac{1}{h}\pi\left(\frac{a_2-\alpha_2}{h}\right)K\left(\frac{a_2-\alpha_2}{h}\right)\\
&\quad\times\left\{\int_0^{a_2}\left[\int_{-\infty}^{a_1}\frac{1}{h}\pi'\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da\right]da_1\right\}da_2d\alpha_1d\alpha_2.
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{J}_h &= \int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\int_0^1 k_h(a_2;\alpha_2)\,a_2\mathbb{E}\left[K_h(A;\alpha_1)\right]'da_2\\
&= \int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\int_0^1 \frac{1}{h}\pi\left(\frac{a_2-\alpha_2}{h}\right)K\left(\frac{a_2-\alpha_2}{h}\right)a_2\\
&\quad\times\left\{\int_0^1\left[\int_{-\infty}^{a_1}\frac{1}{h}\pi'\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da\right]da_1\right\}da_2d\alpha_2d\alpha_1\\
&= \int_0^1 g(\alpha_2)\left[\int_0^1\frac{1}{h}\pi\left(\frac{a_2-\alpha_2}{h}\right)K\left(\frac{a_2-\alpha_2}{h}\right)a_2da_2\right]d\alpha_2\\
&\quad\times\int_0^1 f(\alpha_1)\int_0^1\left[\int_{-\infty}^{a_1}\frac{1}{h}\pi'\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da\right]da_1d\alpha_1.
\end{aligned}
$$

Consider first $\mathcal{J}_h$. The change of variable $a_2 = \alpha_2 + ht$ and the definition of $\Omega_h(\alpha_2)$ give

$$\begin{aligned}
&\int_0^1 g(\alpha_2)\left[\int_0^1 \frac{1}{h}\pi\left(\frac{a_2-\alpha_2}{h}\right)K\left(\frac{a_2-\alpha_2}{h}\right)a_2 da_2\right]d\alpha_2 \\
&\quad = \int_0^1 g(\alpha_2)\left[\int_{-\frac{\alpha_2}{h}}^{\frac{1-\alpha_2}{h}} (\alpha_2+ht)\,\pi(t)\,K(t)\,dt\right]d\alpha_2 \\
&\quad = \int_0^1 \alpha_2 g(\alpha_2)\,\Omega_h(\alpha_2)\,s_0 d\alpha_2 + h\int_0^1 g(\alpha_2)\,\Omega_h(\alpha_2)\,s_1 d\alpha_2.
\end{aligned}$$

For the second item in $\mathcal{J}_h$, integrating by parts gives

$$\begin{aligned}
&\int_0^1\left[\int_{-\infty}^{a_1}\frac{1}{h}\pi'\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da\right]da_1 \\
&\quad = \int_{-\infty}^1 \frac{1}{h}\pi'\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da - \int_0^1 \frac{1}{h}\pi'\left(\frac{a_1-\alpha_1}{h}\right)K\left(\frac{a_1-\alpha_1}{h}\right)a_1 da_1.
\end{aligned}$$

This gives

$$\begin{aligned}
&\int_0^1 f(\alpha_1)\int_0^1\left[\int_{-\infty}^{a_1}\frac{1}{h}\pi\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da\right]da_1 d\alpha_1 \\
&\quad = \int_0^1 f(\alpha_1)\left[\int_{-\infty}^0 + \int_0^1\left\{\frac{1}{h}\pi\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)\right\}da\right]d\alpha_1 \\
&\qquad\quad - \int_0^1 f(\alpha_1)\left[\int_0^1\frac{1}{h}\pi\left(\frac{a_1-\alpha_1}{h}\right)K\left(\frac{a_1-\alpha_1}{h}\right)a_1 da_1\right]d\alpha_1 \\
&\quad = \int_0^1 f(\alpha_1)\left[\int_{-\infty}^{-\frac{\alpha_1}{h}} + \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}}\pi(t)\,K(t)\,dt\right]d\alpha_1 \\
&\qquad - \int_0^1 f(\alpha_1)\left[\int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}}\pi(t)\,K(t)\,(\alpha_1+ht)\,dt\right]d\alpha_1 \\
&\quad = \int_0^1 f(\alpha_1)(1-\alpha_1)\,\Omega_h(\alpha_1)\,s_0 d\alpha_1 - h\int_0^1 f(\alpha_1)\,\Omega_h(\alpha_1)\,s_1 d\alpha_1 \\
&\qquad\quad + \int_0^1 f(\alpha_1)\left[\int_{-\infty}^{-\frac{\alpha_1}{h}}\pi(t)\,K(t)\,dt\right]d\alpha_1.
\end{aligned}$$

Hence

$$\begin{aligned}
\mathcal{J}_h = &\left[\int_0^1 \alpha g\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] s_0 s_0' \left[\int_0^1 f\left(\alpha\right)\left(1-\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] \\
&+ h\left[\int_0^1 g\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] s_1 s_0' \left[\int_0^1 f\left(\alpha\right)\left(1-\alpha\right)\Omega_h\left(\alpha\right)s_0 d\alpha\right] \\
&- h\left[\int_0^1 \alpha g\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] s_0 s_1' \left[\int_0^1 f\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] \\
&- h^2\left[\int_0^1 g\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] s_1 s_1' \left[\int_0^1 f\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] \\
&+ \left[\int_0^1 g\left(\alpha\right)\Omega_h\left(\alpha\right)\left[\alpha s_0 + h s_1\right]d\alpha\right]\left[\int_0^1 f\left(\alpha\right)\left[\int_{-\infty}^{-\frac{\alpha}{h}} \pi'\left(t\right)K\left(t\right)dt\right]d\alpha\right].
\end{aligned}$$

Consider now $\mathcal{I}_h$, which satisfies

$$\begin{aligned}
\mathcal{I}_h &\overset{a_2=\alpha_2+ht_2}{=} \int_0^1\int_0^1 f\left(\alpha_1\right)g\left(\alpha_2\right)\int_{-\frac{\alpha_2}{h}}^{\frac{1-\alpha_2}{h}} \pi\left(t_2\right)K\left(t_2\right) \\
&\quad \times \left\{\int_0^{\alpha_2+ht_2}\left[\int_{-\infty}^{a_1}\frac{1}{h}\pi'\left(\frac{a-\alpha_1}{h}\right)K\left(\frac{a-\alpha_1}{h}\right)da\right]da_1\right\}dt_2 d\alpha_1 d\alpha_2 \\
&\overset{a=\alpha_1+ht}{=} \int_0^1\int_0^1 f\left(\alpha_1\right)g\left(\alpha_2\right)\int_{-\frac{\alpha_2}{h}}^{\frac{1-\alpha_2}{h}} \pi\left(t_2\right)K\left(t_2\right) \\
&\quad \times \left\{\int_0^{\alpha_2+ht_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}} \pi'\left(t\right)K\left(t\right)dt\right]da_1\right\}dt_2 d\alpha_1 d\alpha_2.
\end{aligned}$$

Observe

$$\int_0^{\alpha_2+ht_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right]da_1=\int_0^{\alpha_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right]da_1$$
$$+\int_{\alpha_2}^{\alpha_2+ht_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right]d\left[a_1-\alpha_2-ht_2\right]$$
$$=\int_0^{\alpha_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right]da_1+ht_2\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}}\pi'(t)\,K(t)\,dt$$
$$-\int_{\alpha_2}^{\alpha_2+ht_2}(a_1-\alpha_2-ht_2)\,\frac{1}{h}\pi'\left(\frac{a_1-\alpha_1}{h}\right)K\left(\frac{a_1-\alpha_1}{h}\right)da_1$$
$$=\int_0^{\alpha_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right]da_1+ht_2\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}}\pi'(t)\,K(t)\,dt$$
$$-h^2t_2^2\int_0^1(1-u)\,\frac{1}{h}\pi'\left(\frac{\alpha_2+ht_2u-\alpha_1}{h}\right)K\left(\frac{\alpha_2+ht_2u-\alpha_1}{h}\right)du$$

It follows that $\mathcal{I}_h=\mathcal{I}_0+h\mathcal{I}_1-h^2\mathcal{I}_2$ with

$$\mathcal{I}_0=\int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\,\Omega_h(\alpha_2)\,s_0\left\{\int_0^{\alpha_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right]da_1\right\}d\alpha_1d\alpha_2,$$
$$\mathcal{I}_1=\int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\,\Omega_h(\alpha_2)\,s_1\left\{\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right\}d\alpha_1d\alpha_2,$$
$$\mathcal{I}_2=\int_0^1\int_0^1 f(\alpha_1)\,g(\alpha_2)\int_{-\frac{\alpha_2}{h}}^{\frac{1-\alpha_2}{h}}t^2\pi(t)\,K(t)$$
$$\times\left\{\int_0^1(1-u)\,\frac{1}{h}\pi'\left(\frac{\alpha_2+htu-\alpha_1}{h}\right)K\left(\frac{\alpha_2+htu-\alpha_1}{h}\right)du\right\}dtd\alpha_1d\alpha_2.$$

Consider first $\mathcal{I}_0$. Integrating by parts gives

$$\mathcal{I}_0=\int_0^1 f(\alpha_1)\left\{\int_0^1\left(\int_0^{\alpha_2}\left[\int_{-\infty}^{\frac{a_1-\alpha_1}{h}}\pi(t)\,K(t)\,dt\right]da_1\right)d\left[-\int_{\alpha_2}^1 g(a_2)\,\Omega_h(a_2)\,s_0da_2\right]'\right\}'d\alpha_1$$
$$=\int_0^1 f(\alpha_1)\left\{\int_0^1\left(\int_{\alpha_2}^1 g(a_2)\,\Omega_h(a_2)\,s_0da_2\right)\left(\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}}\pi'(t)\,K(t)\,dt\right)d\alpha_2\right\}d\alpha_1.$$

Integrating again by parts gives

$$\begin{aligned}
&\int_0^1 \left(\int_{\alpha_2}^1 g(a_2)\,\Omega_h(a_2)\,s_0 da_2\right)\left(\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}} \pi'(t)\,K(t)\,dt\right) d\alpha_2 \\
&\quad = \int_0^1 \left[\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}} \pi(t)\,K(t)\,dt \left\{-d\int_{\alpha_2}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right\}'\right]' \\
&\quad = \int_0^1 \left[\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da\right] d\alpha \times \int_{-\infty}^{-\frac{\alpha_1}{h}} \pi'(t)\,K(t)\,dt \\
&\qquad + \int_0^1 \left[\int_{\alpha_2}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right] \frac{1}{h}\pi'\left(\frac{\alpha_2-\alpha_1}{h}\right) K\left(\frac{\alpha_2-\alpha_1}{h}\right) d\alpha_2.
\end{aligned}$$

It holds, for the second item

$$\begin{aligned}
&\int_0^1 \left[\int_{\alpha_2}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right] \frac{1}{h}\pi'\left(\frac{\alpha_2-\alpha_1}{h}\right) K\left(\frac{\alpha_2-\alpha_1}{h}\right) d\alpha_2 \\
&\quad = \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} \left[\int_{\alpha_1+ht}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right] \pi'(t)\,K(t)\,dt \\
&\quad = \left[\int_{\alpha_1}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right] s_0'\Omega_h(\alpha_1) - h\left[\int_{\alpha_1}^1 g(a)\,\Omega_h(a)\,s_0 d\alpha\right] s_1'\Omega_h(\alpha_1) \\
&\qquad + \frac{h^2}{2} g(\alpha_1)\,\Omega_h(\alpha_1)\,s_0 s_2'\Omega_h(\alpha_1) + o\left(h^2\right),
\end{aligned}$$

where the $o(h^2)$ is uniform over $[h, 1-h]$ and is $O(h^2)$ uniformly over $[0, h]$ and $[1-h, 1]$ under the smoothness assumptions for $f(\cdot)$ and $g(\cdot)$, in which case it contributes for $o(h^2)$ when integrated out of $\alpha_1$. Note that

$$\begin{aligned}
&\int_0^1 \left[\int_{\alpha_1}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right] s_0'\Omega_h(\alpha_1)\,f(\alpha_1)\,d\alpha_1 \\
&\quad = \int_0^1 \left[\int_{\alpha_1}^1\int_{\alpha}^1 g(a)\,\Omega_h(a)\,s_0 da d\alpha\right] d\left[\int_0^{\alpha_1} s_0'\Omega_h(a)\,f(a)\,da\right] \\
&\quad = \int_0^1 \left[\int_{\alpha_1}^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_0' \left[\int_0^{\alpha_1} \Omega_h(a)\,f(a)\,da\right] d\alpha_1.
\end{aligned}$$

This gives, since $\int_0^1 \left[\int_\alpha^1 g\left(a\right)\Omega_h\left(a\right)da\right]d\alpha = \int_0^1 \alpha g\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha$

$$\begin{aligned}
\mathcal{I}_0 &= \int_0^1 \left[\int_\alpha^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_0 s_0' \left[\int_0^\alpha \Omega_h\left(a\right) f\left(a\right)da\right]d\alpha \\
&- h\int_0^1 \left[\int_\alpha^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_0 s_1' \Omega_h\left(\alpha\right) f\left(\alpha\right)d\alpha \\
&+ \frac{h^2}{2}\int_0^1 f\left(\alpha\right) g\left(\alpha\right)\Omega_h\left(\alpha\right) s_0 s_2' \Omega_h\left(\alpha\right)d\alpha + o\left(h^2\right) \\
&+ \left[\int_0^1 \alpha g\left(\alpha\right)\Omega_h\left(\alpha\right)d\alpha\right] s_0 \left[\int_0^1 f\left(\alpha_1\right)\left[\int_{-\infty}^{-\frac{\alpha_1}{h}} \pi'\left(t\right)K\left(t\right)dt\right]d\alpha_1\right]
\end{aligned}$$

Consider now $\mathcal{I}_1$. Integrating by parts gives

$$\begin{aligned}
\mathcal{I}_1 &= \left[\int_0^1 f\left(\alpha_1\right)\left\{\int_0^1 \left[\int_{-\infty}^{\frac{\alpha_2-\alpha_1}{h}} \pi\left(t\right)K\left(t\right)dt\right] d\left[-\int_{\alpha_2}^1 g\left(a\right) s_1' \Omega_h\left(a\right)da\right]\right\}d\alpha_1\right]' \\
&= \left[\int_0^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_1 \int_0^1 f\left(\alpha_1\right)\left[\int_{-\infty}^{-\frac{\alpha_1}{h}} \pi'\left(t\right)K\left(t\right)dt\right]d\alpha_1 \\
&+ \int_0^1 f\left(\alpha_1\right)\left\{\int_0^1 \left[\int_{\alpha_2}^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_1 \frac{1}{h}\pi'\left(\frac{\alpha_2-\alpha_1}{h}\right)K\left(\frac{\alpha_2-\alpha_1}{h}\right)d\alpha_2\right\}d\alpha_1
\end{aligned}$$

with

$$\begin{aligned}
&\int_0^1 f\left(\alpha_1\right)\left\{\int_0^1 \left[\int_{\alpha_2}^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_1 \frac{1}{h}\pi'\left(\frac{\alpha_2-\alpha_1}{h}\right)K\left(\frac{\alpha_2-\alpha_1}{h}\right)d\alpha_2\right\}d\alpha_1 \\
&= \int_0^1 f\left(\alpha_1\right)\left\{\int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} \left[\int_{\alpha_1+ht}^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_1 \pi'\left(t\right)K\left(t\right)dt\right\}d\alpha_1 \\
&= \int_0^1 f\left(\alpha_1\right)\left[\int_{\alpha_1}^1 g\left(a\right)\Omega_h\left(a\right)da\right] s_1 s_0' \Omega_h\left(\alpha_1\right)d\alpha_1 \\
&\quad - h\int_0^1 f\left(\alpha_1\right) g\left(\alpha_1\right)\Omega_h\left(\alpha_1\right) s_1 s_1' \Omega_h\left(\alpha_1\right)d\alpha_1 + o\left(h\right).
\end{aligned}$$

Hence

$$
\begin{aligned}
\mathcal{I}_1 = & \int_0^1 f(\alpha) \left[ \int_\alpha^1 g(a)\, \Omega_h(a)\, da \right] s_1 s_0' \Omega_h(\alpha)\, d\alpha \\
& - h \int_0^1 f(\alpha)\, g(\alpha)\, \Omega_h(\alpha)\, s_1 s_1' \Omega_h(\alpha)\, d\alpha + o(h) \\
& + \left[ \int_0^1 g(a)\, \Omega_h(a)\, da \right] s_1 \left[ \int_0^1 f(\alpha) \left[ \int_{-\infty}^{-\frac{\alpha}{h}} \pi'(t)\, K(t)\, dt \right] d\alpha \right].
\end{aligned}
$$

For $\mathcal{I}_2$, the change of variable $\alpha_2 = \alpha_1 + h\tau$, Assumption H and the conditions on $f(\cdot)$

and $g(\cdot)$ give

$$
\begin{aligned}
\mathcal{I}_2 &= \int_0^1 f(\alpha_1) \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} g(\alpha_1 + h\tau) \int_{-\frac{\alpha_1}{h}-\tau}^{\frac{1-\alpha_1}{h}-\tau} t^2 \pi(t) K(t) \\
&\qquad \times \left\{ \int_0^1 (1-u) \pi'(tu+\tau) K(tu+\tau) du \right\} dt d\tau d\alpha_1 \\
&= \int_0^1 f(\alpha_1) \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} g(\alpha_1) \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} t^2 \pi(t) K(t) \\
&\qquad \times \left\{ \int_0^1 (1-u) \pi'(tu+\tau) K(tu+\tau) du \right\} dt d\tau d\alpha_1 + o(1) \\
&= \int_0^1 f(\alpha_1) g(\alpha_1) \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} t^2 \pi(t) K(t) \\
&\times \left\{ \int_0^1 (1-u) \left[ \int_0^1 \frac{1}{h} \pi' \left( \frac{\alpha_2 + htu - \alpha_1}{h} \right) K \left( \frac{\alpha_2 + htu - \alpha_1}{h} \right) d\alpha_2 \right] du \right\} dt d\alpha_1 + o(1) \\
&= \int_0^1 f(\alpha_1) g(\alpha_1) \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} t^2 \pi(t) K(t) \\
&\qquad \times \left\{ \int_0^1 (1-u) \left[ \int_{-\frac{\alpha_1}{h}+tu}^{\frac{1-\alpha_1}{h}+tu} \pi'(\tau) K(\tau) d\tau \right] du \right\} dt d\alpha_1 + o(1) \\
&= \int_0^1 f(\alpha_1) g(\alpha_1) \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} t^2 \pi(t) K(t) \\
&\qquad \left\{ \int_0^1 (1-u) \left[ \int_{-\frac{\alpha_1}{h}}^{\frac{1-\alpha_1}{h}} \pi'(\tau) K(\tau) d\tau \right] du \right\} dt d\alpha_1 + o(1) \\
&= \frac{1}{2} \int_0^1 f(\alpha) g(\alpha) \Omega_h(\alpha) s_2 s_0' \Omega_h(\alpha) d\alpha + o(1).
\end{aligned}
$$

Now, $\mathcal{I}_h = \mathcal{I}_0 + h\mathcal{I}_1 - h^2\mathcal{I}_2$ and the expressions of $\mathcal{I}_0$, $\mathcal{I}_1$ and $\mathcal{I}_2$ give

$$\begin{aligned}
\mathcal{I}_h = \int_0^1 & \left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_0' \left[\int_0^\alpha \Omega_h(a)\,f(a)\,da\right] d\alpha \\
& + h\int_0^1 f(\alpha)\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right]\left[s_1 s_0' - s_0 s_1'\right]\Omega_h(\alpha)\,d\alpha \\
& - h^2\int_0^1 f(\alpha)\,g(\alpha)\,\Omega_h(\alpha)\,s_1 s_1' \Omega_h(\alpha)\,d\alpha \\
& + \frac{h^2}{2}\int_0^1 f(\alpha)\,g(\alpha)\,\Omega_h(\alpha)\left[s_0 s_2' + s_2 s_0'\right]\Omega_h(\alpha)\,d\alpha + o\left(h^2\right) \\
+ & \left[\int_0^1 g(\alpha)\,\Omega_h(\alpha)\left[\alpha s_0 + h s_1\right] d\alpha\right]\left[\int_0^1 f(\alpha_1)\left[\int_{-\infty}^{-\frac{\alpha_1}{h}} \pi'(t)\,K(t)\,dt\right] d\alpha_1\right]
\end{aligned}$$

We now prepare to compute the expansion of $\mathcal{J}_h - \mathcal{I}_h$. Observe $\int_0^1\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] d\alpha = \int_0^1 \alpha g(\alpha)\,\Omega_h(\alpha)\,d\alpha$, so that

$$\begin{aligned}
& \left[\int_0^1 \alpha g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_0 s_0' \left[\int_0^1 f(\alpha)(1-\alpha)\,\Omega_h(\alpha)\,d\alpha\right] \\
& \qquad - \int_0^1\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_0' \left[\int_0^\alpha \Omega_h(a)\,f(a)\,da\right] d\alpha \\
& = -\left[\int_0^1 \alpha g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_0 s_0' \left[\int_0^1 \alpha f(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] \\
& \qquad + \left[\int_0^1 \alpha g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_0 s_0' \left[\int_0^1 f(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] \\
& \qquad - \int_0^1\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_0' \left[\int_0^1 \Omega_h(a)\,f(a)\,da\right] d\alpha \\
& \qquad + \int_0^1\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_0' \left[\int_\alpha^1 \Omega_h(a)\,f(a)\,da\right] d\alpha \\
& = \int_0^1\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_0' \left[\int_\alpha^1 \Omega_h(a)\,f(a)\,da\right] d\alpha \\
& - \left[\int_0^1 \alpha g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_0 s_0' \left[\int_0^1 f(\alpha)\,\alpha\Omega_h(\alpha)\,d\alpha\right], \\
& = \operatorname{Cov}\left(\int_A^1 g(a)\,\Omega_h(a)\,s_0 da, \int_A^1 f(a)\,\Omega_h(a)\,s_0 da\right).
\end{aligned}$$

Similarly, $\int_0^1\left[\int_0^\alpha f(\alpha)\,\Omega_h(a)\,da\right] d\alpha = \int_0^1 f(\alpha)(1-\alpha)\,\Omega_h(\alpha)\,d\alpha$ gives, after an integration by parts,

$$\begin{aligned}
&\left[\int_0^1 g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_1 s_0' \left[\int_0^1 f(\alpha)(1-\alpha)\,\Omega_h(\alpha)\,d\alpha\right] \\
&\qquad - \int_0^1 f(\alpha)\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_1 s_0' \Omega_h(\alpha)\,d\alpha \\
&\quad = \left[\int_0^1 g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_1 s_0' \left[\int_0^1 \left(\int_0^\alpha \Omega_h(a)\,f(a)\,da\right) d\alpha\right] \\
&\qquad - \int_0^1 g(\alpha)\,\Omega_h(\alpha)\,s_1 s_0' \left[\int_0^\alpha \Omega_h(a)\,f(a)\,da\right] d\alpha \\
&\quad = -\operatorname{Cov}\left(g(A)\,\Omega_h(A)\,s_1, \left[\int_0^A f(a)\,\Omega_h(a)\,da\right] s_0\right) \\
&\quad = \operatorname{Cov}\left(g(A)\,\Omega_h(A)\,s_1, \left[\int_A^1 f(a)\,\Omega_h(a)\,da\right] s_0\right), \\
&\int_0^1 f(\alpha)\left[\int_\alpha^1 g(a)\,\Omega_h(a)\,da\right] s_0 s_1' \Omega_h(\alpha)\,d\alpha \\
&\qquad - \left[\int_0^1 \alpha g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] s_0 s_1' \left[\int_0^1 f(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] \\
&\quad = \operatorname{Cov}\left(\left[\int_A^1 g(a)\,\Omega_h(a)\,da\right] s_0, f(A)\,\Omega_h(A)\,s_1\right),
\end{aligned}$$

and, for any conformable $u$ and $v$,

$$\begin{aligned}
&\int_0^1 f(\alpha)\,g(\alpha)\,\Omega_h(\alpha)\,[uv']\,\Omega_h(\alpha)\,d\alpha \\
&\qquad - \left[\int_0^1 g(\alpha)\,\Omega_h(\alpha)\,d\alpha\right][uv']\left[\int_0^1 f(\alpha)\,\Omega_h(\alpha)\,d\alpha\right] \\
&\quad = \operatorname{Cov}\left(g(A)\,\Omega_h(A)\,u, f(A)\,\Omega_h(A)\,v\right).
\end{aligned}$$

Collecting these items gives the expansion of $\mathcal{C}_h$ stated in the Lemma. □

# References

[1] Guerre, E. & C. Sabbah (2012). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory* **28**, 87–129.

[2] Guerre, E. & C. Sabbah (2014). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. http://arxiv.org/pdf/1105.5038.pdf

[3] Massart, P. (2007). *Concentration inequalities and model selection.* Lectures Notes in Mathematics 1986. Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003, Jean Picard (ed.). Springer-Verlag.

[4] van de Geer, S. (1999). *Empirical processes in M-estimation.* Cambridge University Press.